

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2004, 21:00:05 ; Search time 16233 Seconds
(without alignments)
7008.899 Million cell updates/sec

Title: US-10-003-405-1

Perfect score: 2625
Sequence: 1 atgagctgagcgtctctcac.....tcaaaaaaaaaaaaaa 2625

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: GenBdb1.*
2: gb_da.*
3: gb_hcg.*
4: gb_in.*
5: gb_om.*
6: gb_ov.*
7: gb_pat.*
8: gb_ph.*
9: gb_pl.*
10: gb_pr.*
11: gb_ro.*
12: gb_sts.*
13: gb_sy.*
14: gb_un.*
15: gb_vi.*
16: em_ba.*
17: em_fun.*
18: em_hum.*
19: em_in.*
20: em_mu.*
21: em_om.*
22: em_or.*
23: em_ov.*
24: em_pat.*
25: em_ph.*
26: em_pl.*
27: em_ro.*
28: em_sts.*
29: em_un.*
30: em_vi.*
31: em_hcg_hum.*
32: em_hcg_inv.*
33: em_hcg_other.*
34: em_hcg_mus.*
35: em_hcg_pln.*
36: em_hcg_rtd.*
37: em_hcg_mam.*
38: em_hcg_vrt.*
39: em_hcg_hum.*
40: em_hcg_mus.*
41: em_hcg_other.*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2624	100.0	2625	8 U73588	U73588 Gossypium h
2	1689.6	64.4	2661	8 AB022092	AB022092 Citrus un
3	1615.6	61.5	2652	8 VIRUS81	D10266 Vigna radia
4	1602.6	61.1	2842	8 AF030231	AF030231 Glycine m
5	1577.2	60.1	2738	8 AF315375	AF315375 Phaseolus
6	1559	59.4	2728	8 MTR131943	MTR131943 Medicago
7	1556	59.3	2652	8 PSA012080	A012080 Pisum sat
8	1556	59.3	2760	8 AF049487	AF049487 Medicago
9	1555.6	59.3	2770	8 AB018561	AB018561 Citrullus
10	1548.6	59.0	2749	8 AF079851	AF079851 Pisum sat
11	1539	58.6	2647	8 VFSUCS	X69773 V. faba mRNA
12	1539	58.6	2665	8 VFAUDPGFTA	M97551 Vicia faba
13	1499.4	57.1	2906	6 BD235993	BD235993 Materials
14	1497.8	57.1	2913	6 BD235993	BD235993 Materials
15	1497.8	57.1	3103	6 BD262165	BD262165 Compositi
16	1497.8	57.1	3103	6 AR360950	AR360950 Sequence
17	1486.6	56.6	2427	6 AX506925	AX506925 Sequence
18	1486.6	56.6	2427	6 AX651931	AX651931 Sequence
19	1477.6	56.3	2783	8 AGSUS1	X92378 A. glutinosa
20	1475.6	56.2	2757	8 PSB311496	AB311496 Pisum sat
21	1454	55.4	2725	8 TOMSSF	L19762 Lycopersico
22	1452.4	55.3	2711	8 POTSSYN	M18745 Potato sucro
23	1447.8	55.2	2708	8 LES011319	A011319 Lycopersi
24	1447.6	55.1	2701	8 AY205084	AY205084 Solanum t
25	1442.8	55.0	2429	8 STU57575	AY57575 Solanum t
26	1405.4	53.5	2838	8 AY457173	AY457173 Beta vulg
27	1396.4	53.2	2554	8 CRSUCSYN	X62504 C. rudrum MK
28	1380.8	52.6	2866	8 DCRNAS	X75332 D. carota (N
29	1316	50.1	2717	8 AF263384	AF263384 Saccharum
30	1315.4	50.1	2910	8 AF412037	AF412037 Bambusa o
31	1313.6	50.1	2771	8 AK100306	AK100306 Oryza sat
32	1310.2	49.9	2690	8 AF412038	AF412038 Bambusa o
33	1307	49.8	2774	8 AF412039	AF412039 Bambusa o
34	1301.8	49.6	2890	8 AF412036	AF412036 Bambusa o
35	1300.6	49.5	2494	6 E55472	E55472 Method for
36	1300.6	49.5	3389	8 AK100546	AK100546 Oryza sat
37	1299	49.5	2702	8 AK098923	AK098923 Oryza sat
38	1287.4	49.4	2627	8 OSSUPHSY	Z15078 O. sativa MK
39	1286.2	49.4	2885	8 TGSUCSY21	X56939 T. geseenerian
40	1285.8	49.4	2781	8 AK072074	AK072074 Oryza sat
41	1287.6	49.1	2450	6 AX654790	AX654790 Sequence
42	1287.6	49.1	2451	6 AX755420	AX755420 Sequence
43	1287.6	49.1	2451	6 AX755773	AX755773 Sequence
44	1287.6	49.1	2910	8 AK100334	AK100334 Oryza sat
45	1285.4	49.0	2759	8 TGSUCSYN1	X56938 T. geseenerian

ALIGNMENTS

RESULT 1
LOCUS U73588 2625 bp mRNA linear PLN 04-MAY-1999
DEFINITION Gossypium hirsutum sucrose synthase mRNA, complete cds.
ACCESSION U73588
VERSION U73588.2 GI:4733945
KEYWORDS
SOURCE Gossypium hirsutum (upland cotton)
ORGANISM Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE
AUTHORS Perez-Grau, L. and Delmer, D.
TITLE Direct Submission

JOURNAL Submitted (07-OCT-1996) Calgene, Inc., 1920 Fifth Street, Davis, CA 95616, USA
COMMENT On May 4, 1999 this sequence version replaced gi:4098126.
FEATURES Location/Qualifiers
source 1. 2625
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/strain="Acala S-2"
/db_xref="taxon:3635"
/feature_type="cotton fiber"
1. 2421
/codon_start=1
/product="sucrose synthase"
/protein_id="AAD28641.1"
/db_xref="gi:4733946"
/translation="MAERALTRVHSLRLEDETLAHRNELIALLSRIBGKGLIHH
HQLLEBAPALPEBRKLANGAAPREULKASGEALVLPWVALAPRPBGWERYRVV
HALVBEITVAEYLFHFKELVDSSNGNFVLEIDPEPNSSFPRLPTSKISGNCVEL
NRHLSAKLFHDKESMHPLEFLRVHCHGKMKMLNDRIQLNLALQHLRKAEEYGLT
PPEPCAEFEHRFOEIGLERGMGTAEVLEMIQLDLLEATDPCLEKFLGRIPV
FNVLTFPHGYFADNVLYGPDGQVYVILDOVRALENEMLLIKOGLNTPRIIL
ITLLPDAVGTGGRLEKVGTEHSIDLVPRTKEGIVKMTSREKWPYLETY
EDVAHESKEIAGHPDLIGKXSDNIVASLLAKIGVTOCTTAALKEKYPDSDIY
WKJLEDKHRSQFTADLPAMHTDPIITSTFOEIRASKDTVGQYESHPTAFLELYR
VHGIIDVDFPKFNIVSGADMEIFYPTIEERRLKHPIEDLLYTKVENEHLCVL
NDRKRPILFTMRPLDRVNLTVGLVWCGKNEKRLRELNLVVGGRKRESNLEKKE
MKRMELIKYNLNGOFKWLSSQNMRLNVELRYICDTKGAFOPLALYEFGLVVE
AMTGCLPTFATCNGSPAEIIVHSGFIDIPHDADAILVDFEKKKDPSPWMDI
SOGGLKRIEERKTYTKIYSERLLTLTGYYGPFKHYVSNLERRSRYLLEMPALKRKLA
ESVPLAE"

CDS

1. 2421
/codon_start=1
/product="sucrose synthase"
/protein_id="AAD28641.1"
/db_xref="gi:4733946"
/translation="MAERALTRVHSLRLEDETLAHRNELIALLSRIBGKGLIHH
HQLLEBAPALPEBRKLANGAAPREULKASGEALVLPWVALAPRPBGWERYRVV
HALVBEITVAEYLFHFKELVDSSNGNFVLEIDPEPNSSFPRLPTSKISGNCVEL
NRHLSAKLFHDKESMHPLEFLRVHCHGKMKMLNDRIQLNLALQHLRKAEEYGLT
PPEPCAEFEHRFOEIGLERGMGTAEVLEMIQLDLLEATDPCLEKFLGRIPV
FNVLTFPHGYFADNVLYGPDGQVYVILDOVRALENEMLLIKOGLNTPRIIL
ITLLPDAVGTGGRLEKVGTEHSIDLVPRTKEGIVKMTSREKWPYLETY
EDVAHESKEIAGHPDLIGKXSDNIVASLLAKIGVTOCTTAALKEKYPDSDIY
WKJLEDKHRSQFTADLPAMHTDPIITSTFOEIRASKDTVGQYESHPTAFLELYR
VHGIIDVDFPKFNIVSGADMEIFYPTIEERRLKHPIEDLLYTKVENEHLCVL
NDRKRPILFTMRPLDRVNLTVGLVWCGKNEKRLRELNLVVGGRKRESNLEKKE
MKRMELIKYNLNGOFKWLSSQNMRLNVELRYICDTKGAFOPLALYEFGLVVE
AMTGCLPTFATCNGSPAEIIVHSGFIDIPHDADAILVDFEKKKDPSPWMDI
SOGGLKRIEERKTYTKIYSERLLTLTGYYGPFKHYVSNLERRSRYLLEMPALKRKLA
ESVPLAE"
2608 . 2625

ORIGIN
polya_site

Query Match 100.0%; Score 2624; DB 8; Length 2625;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTGAGCGTGTCTTCACTCGGCTCCAGAGTCTCCGTGAGCGTTTGATGAGACCTT 60
DB 1 ATGGCTGAGCGTGTCTTCACTCGGCTCCAGAGTCTCCGTGAGCGTTTGATGAGACCTT 60
QY 61 CTGTGCTCAGAGAAAGAGATTTTGGCTTGTCTCAGAGATCGAGGCGAAAGAAAAGGA 120
DB 61 CTGTGCTCAGAGAAAGAGATTTTGGCTTGTCTCAGAGATCGAGGCGAAAGAAAAGGA 120
QY 121 ATTGCGAAGCAACATCAATTAATCTAGATTGAAGTATCCCTGAAGAGACAGAAAG 180
DB 121 ATTGCGAAGCAACATCAATTAATCTAGATTGAAGTATCCCTGAAGAGACAGAAAG 180
QY 181 AAGCTCGTAATGTGCAATTTTGAAGATTAAGAGCTAGTCAGAAAGCATGCTGTG 240
DB 181 AAGCTCGTAATGTGCAATTTTGAAGATTAAGAGCTAGTCAGAAAGCATGCTGTG 240
QY 241 CCTGCATGGGTGCACTTGTCTGTCAGAGCGCTGATGTTTGGAGATCAATTAGAGTG 300
DB 241 CCTGCATGGGTGCACTTGTCTGTCAGAGCGCTGATGTTTGGAGATCAATTAGAGTG 300
QY 301 AATGTCAAGCGCCTGTGTTGAGGAATCACTGTGTGAGTATCTCACTTCAAGGAA 360
DB 301 AATGTCAAGCGCCTGTGTTGAGGAATCACTGTGTGAGTATCTCACTTCAAGGAA 360
QY 361 GAGCTGTGTAAGAAAGTTCAATGGAACCTTGTGTAATTGGAATTTGAGACCCCTTC 420
DB 361 GAGCTGTGTAAGAAAGTTCAATGGAACCTTGTGTAATTGGAATTTGAGACCCCTTC 420
QY 421 AACTCATATATCCCGCCCAACTCTTCAAAATTCATTTGATGATGATGATGATGATG 480
DB 421 AACTCATATATCCCGCCCAACTCTTCAAAATTCATTTGATGATGATGATGATGATG 480
QY 481 AATGTCACTTTGGGCAAAATTTGTTCAATGCAAGAGAGATGACCCCTTGTCTGAA 540
DB 481 AATGTCACTTTGGGCAAAATTTGTTCAATGCAAGAGAGATGACCCCTTGTCTGAA 540

QY 541 TTCTCAGAGTCCATGTTCTCAAGAGGCGAAGACATGATGTTGAATGACAGAAATTCAGAAC 600
DB 541 TTCTCAGAGTCCATGTTCTCAAGAGGCGAAGACATGATGTTGAATGACAGAAATTCAGAAC 600
QY 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGCAGAGAGATCTTGTATCCATCTCTCT 660
DB 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGCAGAGAGATCTTGTATCCATCTCTCT 660
QY 661 GAGACACCATGTCGCAATTCGAACACCGGTTCAGAGAAATCGGTTTGAAGAGTTGG 720
DB 661 GAGACACCATGTCGCAATTCGAACACCGGTTCAGAGAAATCGGTTTGAAGAGTTGG 720
QY 721 GGTGACACCGCAGAACGCGTGTCTGAGATGATCACTCTTTTGATCTTCTTGAGGCA 780
DB 721 GGTGACACCGCAGAACGCGTGTCTGAGATGATCACTCTTTTGATCTTCTTGAGGCA 780
QY 781 ACTGATCTTGCACCCCTTGAGAAAGTTCCCTTGGAGAAATCCCATNGGTTCAATGTTGTG 840
DB 781 ACTGATCTTGCACCCCTTGAGAAAGTTCCCTTGGAGAAATCCCATNGGTTCAATGTTGTG 840
QY 841 ATTCTCACTCCCAACGGAATCTTGTCTCAAGACAAATGTTTGGGATATCCCGACACCGGT 900
DB 841 ATTCTCACTCCCAACGGAATCTTGTCTCAAGACAAATGTTTGGGATATCCCGACACCGGT 900
QY 901 GGGCAGGTTGTTATCATCTTGAATCAAGTCCGAGCTTTGAGAAATGAGATGCTCTCCGT 960
DB 901 GGGCAGGTTGTTATCATCTTGAATCAAGTCCGAGCTTTGAGAAATGAGATGCTCTCCGT 960
QY 961 ATTAAGCAAGAGACATCAATCAACCCCTCGAATCTCTATTAATCACTTCTCTCT 1020
DB 961 ATTAAGCAAGAGACATCAATCAACCCCTCGAATCTCTATTAATCACTTCTCTCT 1020
QY 1021 GATGCTGTCCGAAACAACATCGGTCAACGACTTGAGAAATATACGAAACAGACACTCG 1080
DB 1021 GATGCTGTCCGAAACAACATCGGTCAACGACTTGAGAAATATACGAAACAGACACTCG 1080
QY 1081 GATATTTCTGAGTACCTTCAAGACAGAAAGGAAATTTGGAATAATGATCTCAAGA 1140
DB 1081 GATATTTCTGAGTACCTTCAAGACAGAAAGGAAATTTGGAATAATGATCTCAAGA 1140
QY 1141 TTGAAAAGTCTGCGCATCTTGAAGAACTTACAGAGAGATGTCATGAAATCTCC 1200
DB 1141 TTGAAAAGTCTGCGCATCTTGAAGAACTTACAGAGAGATGTCATGAAATCTCC 1200
QY 1201 AAAGAGTTGACGCGCACGCGAGATCTGATCATGSAACAAACAGCGCAATATGCTC 1260
DB 1201 AAAGAGTTGACGCGCACGCGAGATCTGATCATGSAACAAACAGCGCAATATGCTC 1260
QY 1261 GCTCTCTGCTCGACATTAATTAAGTGTCAACAGTGCACATCGCCATGCTTGGAG 1320
DB 1261 GCTCTCTGCTCGACATTAATTAAGTGTCAACAGTGCACATCGCCATGCTTGGAG 1320
QY 1321 AAGCAAAATATCAAGTTGAGATCTATTTGAAGAAAGCTTGAAGCAAAATCAATTC 1380
DB 1321 AAGCAAAATATCAAGTTGAGATCTATTTGAAGAAAGCTTGAAGCAAAATCAATTC 1380
QY 1381 TCTTGCAATTTACAGTGAATCTTCTTGAAGAAACCATATCAGATTCATCATCACAGT 1440
DB 1381 TCTTGCAATTTACAGTGAATCTTCTTGAAGAAACCATATCAGATTCATCATCACAGT 1440
QY 1441 ACTTTCAGAGAAATGAGAGAGAGACATGTTGTCATATACAGAGACACATGCT 1500
DB 1441 ACTTTCAGAGAAATGAGAGAGAGACATGTTGTCATATACAGAGACACATGCT 1500
QY 1501 TTGACCTTCTGCTGCTCAACGCGTGTGATCAATGATGATGATGATGATGATGATG 1560
DB 1501 TTGACCTTCTGCTGCTCAACGCGTGTGATCAATGATGATGATGATGATGATGATG 1560
QY 1561 AACTGTTTCCCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1620
DB 1561 AACTGTTTCCCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1620
QY 1621 AGGTGAAGCATTTCCATCTGAGATGGAAGACCTTCTTTACACAAAGTTGAGATGAA 1680

```

Db 1621 AGGTGAAGCATTTTCATCTGAGATCGAAGACCTTTTACACCAAGTTGAGATGAA 1680
Qy 1661 GAACCTATGTGCTGCAATGACCGCAACAGCCAAATTCGTTCACAAATGCCAAGGCT 1740
Db 1661 GAACCTATGTGCTGCAATGACCGCAACAGCCAAATTCGTTCACAAATGCCAAGGCT 1740
Qy 1741 GATCGTGTCAAGAACTTAACCGGACTCGTCGAGTGGTGGCGGAGAACCCAAAGTTGGCT 1800
Db 1741 GATCGTGTCAAGAACTTAACCGGACTCGTCGAGTGGTGGCGGAGAACCCAAAGTTGGCT 1800
Qy 1801 GAGTTGGCTAACTCTGAGTGTGTAGTGTGTAGCGCAAGAAAGTCTAAAGATTGGAA 1860
Db 1801 GAGTTGGCTAACTCTGAGTGTGTAGTGTGTAGCGCAAGAAAGTCTAAAGATTGGAA 1860
Qy 1861 GAGAAAGCTGAAATGAAAGAAATGTTTGAAGCTGATCGAAGTACACTTGAACGGCCAA 1920
Db 1861 GAGAAAGCTGAAATGAAAGAAATGTTTGAAGCTGATCGAAGTACACTTGAACGGCCAA 1920
Qy 1921 TTCAGATGATATCATCTCAATGAAAGAAATCCGAAATGTTGAACCTTACCGATACATT 1980
Db 1921 TTCAGATGATATCATCTCAATGAAAGAAATCCGAAATGTTGAACCTTACCGATACATT 1980
Qy 1981 TCGCAACAGAAAGGTGCTTTGTACAGCTGATGTAAGACCTTTGAAATGACAGTT 2040
Db 1981 TCGCAACAGAAAGGTGCTTTGTACAGCTGATGTAAGACCTTTGAAATGACAGTT 2040
Qy 2041 GTGAGGCGCAATGACTTGGGCTTTCGCAACATTCGCAACCTGTAACGGTGGACCGCAG 2100
Db 2041 GTGAGGCGCAATGACTTGGGCTTTCGCAACATTCGCAACCTGTAACGGTGGACCGCAG 2100
Qy 2101 ATTATGTCACAGGAAATCTGTTTCAACATGATCCTTACCATGATGATCAAGCTGCT 2160
Db 2101 ATTATGTCACAGGAAATCTGTTTCAACATGATCCTTACCATGATGATCAAGCTGCT 2160
Qy 2161 GACATACTCGTCGATTTCTTGAAGAAAGTGAAGAAAGTCCATCTCACTGGGATTAAGATC 2220
Db 2161 GACATACTCGTCGATTTCTTGAAGAAAGTGAAGAAAGTCCATCTCACTGGGATTAAGATC 2220
Qy 2221 TCCCAAGAGAGCTTGAAGCAATGAGAGAAAGTATACATGAGAGATTTACTCGGAGGA 2280
Db 2221 TCCCAAGAGAGCTTGAAGCAATGAGAGAAAGTATACATGAGAGATTTACTCGGAGGA 2280
Qy 2281 CTATTGACCTCGACAGAGATGATGATTCCTGGAAGCAATGTTTCCAACTTGAAGCGGCT 2340
Db 2281 CTATTGACCTCGACAGAGATGATGATTCCTGGAAGCAATGTTTCCAACTTGAAGCGGCT 2340
Qy 2341 GAGAGTCGTGCTTACCTTGAGATGTTTATGCTCTTAAGTACCGTAACTGGCTGAATCA 2400
Db 2341 GAGAGTCGTGCTTACCTTGAGATGTTTATGCTCTTAAGTACCGTAACTGGCTGAATCA 2400
Qy 2401 GTTCCATTGGCAGAGAGATTAATGAACCTGTTAAATACATTTGAGGCGGCTTTCTTGG 2460
Db 2401 GTTCCATTGGCAGAGAGATTAATGAACCTGTTAAATACATTTGAGGCGGCTTTCTTGG 2460
Qy 2461 AGAATTAATATCTGTTTGTAAATTTCAATGAGAGAGCTCTTGTATTTCAATCTGCT 2520
Db 2461 AGAATTAATATCTGTTTGTAAATTTCAATGAGAGAGCTCTTGTATTTCAATCTGCT 2520
Qy 2521 TTTCCCTTTCTTTTTCGCGCGCATTTGTTGAACATGGGGTGTGCGCCCGCTCAATTC 2580
Db 2521 TTTCCCTTTCTTTTTCGCGCGCATTTGTTGAACATGGGGTGTGCGCCCGCTCAATTC 2580
Qy 2561 AGTTAAATATGATGATCTTTGTTTTCACAAAAAATTTTCAAAAAAATTTTCAAAAAA 2625
Db 2561 AGTTAAATATGATGATCTTTGTTTTCACAAAAAATTTTCAAAAAAATTTTCAAAAAA 2625

```

RESULT 2
 AB022092
 LOCUS
 DEFINITION Citrus unshiu CItSUS1 mRNA for sucrose synthase, complete cds.
 ACCESSION AB022092

```

VERSION AB022092.1 GI:6682842
KEYWORDS sucrose synthase.
SOURCE Citrus unshiu
ORGANISM Citrus unshiu
REFERENCE 1
AUTHORS Komatsu,A., Moriguchi,T., Koyama,K., Omura,M. and Akihama,T.
TITLE Analysis of sucrose synthase genes in citrus suggests different
roles and phylogenetic relationships
JOURNAL J. Exp. Bot. 53 (366), 61-71 (2002)
MEDLINE 21606208
PUBMED 11741042
REFERENCE 2 (bases 1 to 2661)
AUTHORS Komatsu,A.
TITLE Direct Submission
JOURNAL Submitted (22-DEC-1998) Akira Komatsu, National Institute of Crop
Science; 2-1-18, Tsukuba Science City, Ibaraki 305-8518, Japan
(B-mail:akomatsu@affrc.go.jp, Tel:81-298-38-8949,
Fax:81-298-38-8949)
FEATURES
Source
1..2661
/organism="Citrus unshiu"
/mol type="mRNA"
/cultivar="Miyagawa-Wase"
/db xref="taxon:55188"
/tissue type="fruit"
/dev stage="middle developmental stage of fruit juice
sacs"
1..2661
/gene="CItSUS1"
24..2441
/gene="CItSUS1"
/codon_start=1
/product="sucrose synthase"
/protein_id="BA88905.1"
/db xref="GI:6682843"
/translation="MAERLTVRHSRLRELDLTSARHRELLALLSRLEGKGLION
HQLAIEPESISEERKHLTEGAREEVLRATQEAIVLAPRWALAVPRPGVREYLRNV
HALVEELVAEYVHAFREELVDGSGNPFVLEDFEPNAPFPPTLSKISGNEVFL
NRHLASLFDHESMPLLEFLRVHCKGKMMINDRIQINLSIQHLRKAEYLTV
VPEPFESEALARFOEIGLERGMDTARALEMIDLDLLEAPPCLEFLGRIPWY
FNVVLPHGVFADVDVGVDPDGGOVVILDDVRALDEDLARIKOOGDITFOILL
ITRLPBAVGTGQRLKRYCTYSLLVLPFTEGVRKXISREWVYLETYE
DVAVEIKELDGKFDLITGNVSDGNIVASLALAGTVOCTITAHLEKTKPDSDIYR
KNLDKXHFSCQFTADLIAMNHTDPIITSTFOELAGSKDVTGQYESHAFLLPELYRV
VHGIDVDPKPNIVSPGADMSIYFPYEBKRLKSPBIEELLVSDVENKHELVK
DRNKPILETARLDRVKNLGLVEMYKNAKRLRELNVVVGDRRKESKDLERQAEK
KRWVSLIDONKLNQFPWISSOMNRVNGELVYICETKGAFPQALYEAAGLVVEA
MTGSLPTPRAICKGSPATIIYNGKSGNHIDPFGEOAALIVDFEKKKADPSVYDKIS
DGLKRIEERTKTIYSORLLTLTGVGFMVHNSNLRLESRRYLEMFLTKRLAE
SVPLAVE"
2661
/gene="CItSUS1"
/note="10 a nucleotides"
ORIGIN
Query Match 64.4%; Score 1689.6; DB 8; Length 2661;
Best Local Similarity 80.3%; Pred. No. 0;
Matches 1993; Conservative 0; Mismatches 485; Indels 3; Gaps 1;
polyA_site
1 ATGGCTGAGCGTGTCTCATCTGCGTGCACAGTCTCCGTGAGCGCTTGATGAACCTT 60
24 ATGGCCGAAACGTGTTGACCCCGGTCACAGTCTCCGCGAGCGCTTGGATGAACCTTC 83
61 CTGTGTCACAGAGAGATTTTGGCTGCTCAAGATCGAGGCGAAAGAAAGAA 120
84 TCTGCCACAGAGATGAATTTGGCCCTTTGTGCAAGATGAGGCGCAAGAGA 143
121 ATCTGCAACCAATCAATTAATTTCTAGAGTTGAAGCTATCCTTGAAGACAAG 180
144 ATCTTGCAAAACATCAATTAATTTCTAGAGTTGAATTTCTGAGGAAAAAGAA 203

```

```
QY 181 AAGCTCGTAATGTCATTTTTTGAAGTATGAAAGCTACTCAGAAAGCATGCTG 240
DB 204 CATCAATCAGAGGTCATTTGTCGAAGTCTAAGAGCCACCCAGAGCAATTTGTTTG 263
QY 241 CCTCAGAGGTTGCACTGCTGTTCTGTCAGAGGCTGCTGTTGGAGTACATTAAGAGT 300
DB 264 GCTCCGAGGTCAGGTCCTGCGTCGTCAGAGGCCGCTGTCGAGTACATCCGAGTG 323
QY 301 AATGTCAGAGCCTTGTTGTTGAGGAATCACTGTTGTCGAGTATCTCCACTTCAAGGA 360
DB 324 AATGTCATGCTCTGTTGTCGAGAGTCTGTTGAGCCAGTACCTGCACTTCAAGAG 383
QY 361 GAGCTTGTGATGAAAGTTCAATGGAACCTTTGTTGAAATTGAAATTTGAGCCCTTC 420
DB 384 GAATCTGTCATGAGAGGCTCCCAATGGAACCTTGGCTTGAGTTGAACTTTGAGCCATTC 443
QY 421 AACTCATATTCCTCCCGCCCAACTTTTCAAATTCATTTGGTAAATGCTGAGATTCCTA 480
DB 444 AATGCTCTGTTCCCTCGCCCAACTTTTCAAAGTCTATTTGAAATGCTGAGAAATTCCTT 503
QY 481 AATGTCACCTTTGCGGCAAAATGTTTCATGACAAAGAGACATGACCTTTGCTGCA 540
DB 504 AACGCCACCTTTTCTGCGAAGCTTTTCATGACAAAGAGACATGACATCTGCTTGA 563
QY 541 TTCTCAGAGTCATTTGTCAGAAAGGCAAGAACATGATGTTGAAATGACAAATTCAGAAC 600
DB 564 TTTCTCGCGTCACCTGCGCAAGAGGCAAAACATGATGTAATGACAAATTCAGAAC 623
QY 601 TTGAATGCTCTTTCAAACATGTTTTCAGAGAAAGCAAGAGATCTTTGGATCCCTACCTCT 660
DB 624 TTGAATTCCTTTCAAACATGTTTTCAGAGAAAGCAAGAGATCTTGAACCAAGTGTACCA 683
QY 661 GAGACACCATGTCGGAATTCGAACACCGGTTCCAGAAATCGGTTTGAAGAGGTTGG 720
DB 684 GAACGCAATTCCTGAGTTGAGTACCTAAGATTTTCAGAGATTTGGTTTGAAGAGGTTGG 743
QY 721 GGTGACACCGGAGAACCGGTCCTGAGATGATTCACCTCTTTTGGATCTTTTGAAGCA 780
DB 744 GGTGATACCTGTCGAGCGTCGACCTTGAGATGATTCAGCTTTTGGATCTTTTGAAGCA 803
QY 781 ACTGATCTTTCGACCTTTGAGAAAGTCTTTCGAGAAATCCCAATGCTGTTCAATGTTGTC 840
DB 804 CCTATTCGTCGACCTTTCGAGATCTTTCGAGAAATCCCAATGCTGTTCAATGTTGTC 863
QY 841 ATTCTCACTCCCAAGATACCTTCTGCTCAAGACAAATGTTTGGGATATCCGACACCGGT 900
DB 864 ATCTTACTCTCCCAAGATACCTTCTGCTCAAGACAAATGTTTGGGATATCCGACACCGGT 923
QY 901 GGGCAGGTTGTTTACATCTTGAATCAAGTCCGAGCTTTGAGAAATGAGATCTCTCCGT 960
DB 924 GGTGAGGTTGTTTACATCTTGAATCAAGTCTGTCCTTGAAGATGAGATCTCTCCGT 983
QY 961 ATAAAGCAAGAGCTCAATCAATCAACCTTCGAAATCCCTATTTATCTAGACTTCTTCT 1020
DB 984 ATCAAGCAAGAGCTCAATCAATCAACCTTCGAAATCCCTATTTATCTAGACTTCTTCT 1043
QY 1021 GATGCTGTCGGAACCAATGCGGTCACAGCTTGAGAAAGTATACGAAACAGACACTCG 1080
DB 1044 GATGCTGTCGGAACCAATGCGGTCACAGCTTGAGAAAGTATATGAAATCAAGTATTTCA 1103
QY 1081 GATAATCTTTCAGATACCTTTCAGAACAGAAAGGAAATGTTTCGAAATGAGATCTCAGA 1140
DB 1104 GATAATCTTTCAGATACCTTTCAGAACAGAAAGGAAATGTTTCGCAAGTGAATCTCAGA 1163
QY 1141 TTTGAAAAAGTCTGGCCATCTTTCGAAACCTTACACAGAGATGTTTCTCATGAAATCTCC 1200
DB 1164 TTTG---AAGTCTGGCCATCTTTCGAAACCTTACACAGAGATGTTTCTCATGAAATCTCC 1220
QY 1201 AAGAGTTGACGCGACCGCAGATCTGATCATCGGAAACNACAGGACCGCAATATGCTC 1260
DB 1221 AAGAGTTGACGCGACCGCAGATCTGATCATCGGAAACNACAGGATGAGCAATGTTT 1280

QY 1261 GCTCTCTGCTCGCACATAATTAAGTGTACACACAGTGCACATCCGCCATGCTTTGGAG 1320
DB 1281 GCTTCTCTGTTAGCACATAAATAGGTTTACCCAGAGTACTATGCTCATGCTTTGGAG 1340
QY 1321 AAGCAAAATATCAGATTCAGATATCTATTTGGAAGAGCTTGAAGCAATATCATTTTC 1380
DB 1341 AAGCAAAATATCAGATTCAGATATCTATTTGGAAGAGCTTGAAGCAATATCATTTTC 1400
QY 1381 TCTTGCCAAATTTACAGCTGATCTTTTTCAGATGAAACCATACAGATTTTCAATCATCAGT 1440
DB 1401 TCAATGCCAATTTACAGCTGATCTTTATGATGATGATGATGATGATGATGATGATGATG 1460
QY 1441 ACTTTCAGAGAAATTCAGAGAGAGAGACACCTGTTGTCATATGAGAGCACAATGCT 1500
DB 1461 ACTTTCAGAGAAATTCAGAGAGAGAGACACCTGTTGTCATATGAGAGCACAATGCT 1520
QY 1501 TTCACTCTCTGCTGCTCAATCCGTTGTCATGATGATGATGATGATGATGATGATGATG 1560
DB 1521 TTCACTCTCTGCTGCTCAATCCGTTGTCATGATGATGATGATGATGATGATGATGATG 1580
QY 1561 AACATGTTTCCCTCGTGTGTCATATGAGATATACCTTCCCTTACACCGAAGAGAACGG 1620
DB 1581 AACATGTTTCCCTCGTGTGTCATATGAGATATACCTTCCCTTACACCGAAGAGAACGG 1640
QY 1621 AGTTGAAGCATTTCCATCTGATGATGAAACCTTCTTTCACCAAGTTGAGATGAA 1680
DB 1641 AGTTGAAGCATTTCCATCTGATGATGAAACCTTCTTTCACCAAGTTGAGATGAA 1700
QY 1681 GAACACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
DB 1701 GAACACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1760
QY 1741 GATGCTGTCAAGAACTTTCACCGGACTGCTGAGTGTGTCGCAAGAACCCAAAGTTGCT 1800
DB 1761 GATGCTGTCAAGAACTTTCACCGGACTGCTGAGTGTGTCGCAAGAACCCAAAGTTGCT 1820
QY 1801 GAGTTGCTTACCTCGTATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1860
DB 1821 GAGTTGCTTACCTCGTATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1880
QY 1861 GAGAGGCTGAATAATGAAGAAATGTTTGAAGCTGATGCAAGTATCAATGAAAGGCA 1920
DB 1881 GAGAGGCTGAATAATGAAGAAATGTTTGAAGCTGATGCAAGTATCAATGAAAGGCA 1940
QY 1921 TTGAGATGATATCATCTCAATATGAAACAGATCCGAATGTTGAATCTTTCAGATCAT 1980
DB 1941 TTGAGATGATATCTCTCTCAATATGAAACAGATCCGAATGTTGAATCTTTCAGATCAT 2000
QY 1981 TCGCACAGAAAGGCTCTTTCGACAGCTGATGATGATGATGATGATGATGATGATGATG 2040
DB 2001 TCGCACAGAAAGGCTCTTTCGACAGCTGATGATGATGATGATGATGATGATGATGATG 2060
QY 2041 GTGAGGCAATGATCTGCGGTTTTCGCAACTTTCGCAACTTTCGCAACTTTCGCAACTTTC 2100
DB 2061 GTGAGGCAATGATCTGCGGTTTTCGCAACTTTCGCAACTTTCGCAACTTTCGCAACTTTC 2120
QY 2101 ATTATGCTCCAGGAAATCTGTTTTCACATGATCTTTCACATGATCTTTCACATGATCT 2160
DB 2121 ATTATGCTCCAGGAAATCTGTTTTCACATGATCTTTCACATGATCTTTCACATGATCT 2180
QY 2161 GACATCTGCTGATCTTTCGTTGAAAGTGTATGAAAGATCAATCTCACTGAGATGATG 2220
DB 2181 GACATCTGCTGATCTTTCGTTGAAAGTGTATGAAAGATCAATCTCACTGAGATGATG 2240
QY 2221 TCCCAAGAGGCTTTCGAAAGATGAGAGAGATGATGATGATGATGATGATGATGATGATG 2280
DB 2241 TCCCAAGAGGCTTTCGAAAGATGAGAGAGATGATGATGATGATGATGATGATGATGATG 2300
QY 2281 CTATTCACCTTCAGAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
DB 2301 CTATTCACCTTCAGAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 2360
QY 2341 GAGAGTGTGCTTACCTTGAAGATGTTTATGCTTTCATGATGATGATGATGATGATGATG 2400
```



```

|||||
Db 2361 GAGAGCCCTCCATACCTGAAATGTTTATGCTCAAGTACCGTAAGCTGGCTGAATCT 2420
Oy 2401 GTTCATGGGAGAGAGTAATGTAACCTGTTAATAATCACTGGCCGGTTCCTGG 2460
Db 2421 GTTCCTGGCTGGTGAAGTAATGACATTTGAGGAAGCTCTCAAGTGGTGGTCCGAGT 2480
Oy 2461 AGAATAATATTTCTGTTTGTGA 2481
Db 2481 GGAATTTCTCGAGTTGTGGA 2501

RESULT 3
VIRVSS1
LOCUS VIRVSS1 2652 bp mRNA linear PLN 02-FEB-1999
DEFINITION Vigna radiata mRNA for sucrose synthase, complete cds.
ACCESSION D10266
VERSION D10266.1 GI:218332
KEYWORDS D-fructose 2-glucosyltransferase; UDPglucose; sucrose synthase;
vss1.
SOURCE Vigna radiata
ORGANISM Vigna radiata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Vigna.
REFERENCE 1 (bases 1 to 2652)
AUTHORS Arai,M., Mori,H. and Imaseki,H.
TITLE Expression of the gene for sucrose synthase during growth of mung
bean seedlings
JOURNAL Plant Cell Physiol. 33, 503-506 (1992)
COMMENT Submitted (21-Jan-1992) to DBJ by:
Masao Arai
National Institute for Basic Biology
Myodaiji-cho, Okazaki
Japan
Phone: 0564-55-7611
Fax: 0564-53-7400
FEATURES
Source Location/Qualifiers
1..2652
/organism="Vigna radiata"
/mol_type="mRNA"
/db_xref="taxon:157791"
/cclone="VSS1"
/cisue_type="stem"
/dev_stage="seedling"
1..2652
/gene="vss1"
38..2455
/gene="vss1"
/EC_number="2.4.1.13"
/product="sucrose synthase"
/codon_start=1
/protein_id="BAA01108.1"
/db_xref="GI:218333"
/translation="MARDRLRVHSLSRDLDTLSANRETLALSLRSGKSLIQLH
HOVAEPEEIPESRKOITDGAEGVLSLSTDEATVLPVWALARPFGVEYRANV
HALVVEVLPAPAEVLRFEKELVDGSSNGFVLELDEPPTASFPPTLANKSINGVOFL
NRHLASAKLFDKESILPLLEFLRLHSVYKRLMLNDRLQNDALQHLVLRKAEYLGTV
PRTPTSAPEHKFOEIGLERGMNAEVLSEIQLDLLEAPDPCLETFLGRIPMV
FNVLISPHGVFADNVLYGPDGGOVYLIDQVRLLENMLHRIKQGLDIVRIILI
ITRLIPAVGTCQORLEKVGTEHSILRVPRPTENGIVAKWISREWPYLETYTE
DVAMELAKEDQKFDLYVNGSDGNTVASLAKLGVTQCTIAALEKTKIPESDIYM
KLEERHFSQPTADLPFAMNHTDPIITSTQELAGSDQVGVYESHTAFTLPGLYV
VHGIDVDPKFNIVSPGADQTIYPHETESRLTSFHELELLSYSENEEHCVLK
DRSKPIILETARLDLRVKNITGLVEMYGNALRLVNLVVVAGDRRKESKDLERKAM
KQWYSLIETVTLNGOFMISOMRVRNGELRYIADTKGAFVQPAYEARGLIVVA
MTGCLPTPATCNGSPAEIIVYKSGFHIIDPHGRADLVLEFEKAKVDPSSHDKIS
OAGLRIEKTMOIYSORLLITLGVGVKHSNLDRESRYLEMPYALKTKKAE
SVPLAVE"
2652
/gene="vss1"
polya_site
/gene="vss1"

```

```

ORIGIN
Query Match 61.5%; Score 1615.6; DB 8; Length 2652;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 1931; Conservative 0; Mismatches 505; Indels 3; Gaps 1;

Oy 1 ATGGCTGAGCGTGTCTCATCTCGGTCACAGTCCTCCGTGAGCGTTTGATGAGACCCCTT 60
Db 38 ATGGCTACCGATCGTTTGACCCCGTGTTCACAGTCTCCGTGAGAGCTTGATGAACCCCTG 97
Oy 61 CTGTCTCAGAGAACGAGATTTTGGCTTGTCTCAAGATGAGAGGCAAGAAAAGGA 120
Db 98 TCTGCAACAGAGAAAGAAATTTCTGGCCCTTCTGTCAAGATGAAAGGCAAGGGG 157
Oy 121 ATTCTGCAACACCAATCAATATTCTAGAGCTTTAGAGCTATCCCTGAAAGAAACGAAG 180
Db 158 ATTTTGCAACACCAATCAAGTATTCAGATTTGAGAAATCCCGAGAGACAGACAG 217
Oy 181 AAGCTCGCTAATGATGATTTTGTGAAGTATTGAAGCTAGTCAAGAACGATCGTGTG 240
Db 218 AAGCTTACTGATGTGCTTGTGAGAAATTTGAGATCTACTCAAGAACCATAGTTTG 277
Oy 241 CTTCCATGAGTGTGACTTGTCTGTTCGTCAGAGCCCTGTGTGTTGGAGTACATTAGAGTG 300
Db 278 CCACCATGAGTGTGCTGTGGCCGTTGTCAGAGCCGTGTGTGTGGAGTACTTGAGAGTG 337
Oy 301 AATGTTACCGCCCTTGTGTGTGAGAACTCATCTGTGTGATCTTCCTACTTCAAGAA 360
Db 338 AATGTGACCGCTTGTGTGTGAGAGTGTGCAACCTGTGTGATCTTCGCTTCAAGAG 397
Oy 361 GAGCTTGTGATGAGAGTCAATGAAACCTTGTGTGAAATTTGAAATTTGAGCCCTTC 420
Db 398 GAATCTGTGATGAGAGTCTTAATGCAACTTTGTGTGAGTTGAGCTTTGAACCTTT 457
Oy 421 AACTCATATTTCCCGCCCACTCTTCAAAATTCATGCTGTAATGTGTGAGTCTTA 480
Db 458 ACCCATCTCTTCCCGCCCACTCTTCAAAATTCATGCTGTAATGTGTGAGTCTTA 517
Oy 481 AATGTCACCTTTTGGCAAAATTTGTCATGACAGAGAGACATGCACTTTGTCTGAA 540
Db 518 AACGTCACTTTTGGCAAACTTTTCATGACAGAGAGACATGCACTTTTGTGAA 577
Oy 541 TTCTCAGAGTCATTTGTCAGAGGCAAGACATGATGTTGATGACAGAAATTCAGAAC 600
Db 578 TTCTCAGAGTCATGAGGTCAAGAGGAAAGACTTGTGATGATGACAGAAATTCAGAAC 637
Oy 601 TTGAATGCTTTCACATGTTTGTGAGAAAGACAGAGAGATCTTGTGATCCCTACT 660
Db 638 CCGATGCTCTTCAACATGTTCTGAGGAAGCTGAGGATCTTGTGATCCCTACT 697
Oy 661 GAGACACCATGTCCGAAATTTGAAACCGGTTCCAGAAATCGGTTTGGAAAGGTTGG 720
Db 698 GAAACCCCTACTAGAGATTTGAGACACAGATTCAGAGATTTGAGAGAGGTTGG 757
Oy 721 GGTGACACCGAGAGAGGCTGTGAGATGATTCACATCTTTTGGATCTTCTGAGGCA 780
Db 758 GGTGACACCGAGAGGCTGTGAGATGATTCACATCTTCTTGGATCTTCTGAGGCA 817
Oy 781 ACTGATCTTGTGACCTTGTGAGAGTCTTGTGAGAAATCCCATGATGTTCAATGTTGTG 840
Db 818 CCGATCTTGTGACCTTGTGAGATCTTGTGAGAAATCCCATGATGTTTCAATGTTGTG 877
Oy 841 ATTCTCACTCCCAAGGATCTTGTGTCAGAGCAATGTTTGGGATATCCGACACCGGT 900
Db 878 ATTCTTCTCCCAAGGATCTTGTGTCAGAGATGATGTTTGGGATATCCGACACCGGT 937
Oy 901 GGGCAGGTGTTTCACTTGTGATCAAGTCGAGCTTGTGAGATGAGATGATGATCTCTCGT 960
Db 938 GGTAGGTTGTTTCACTTGTGATCAAGTCGAGCTTGTGAGATGAGATGATGATCTCTCGT 997
Oy 961 ATTAAGCAACAGAGACTCAACATCACTGATCTGATCTGATCTGATCTGATCTGATCT 1020
Db 998 ATTAAGCAACAGAGATTTGACATGTTCTCTGATCTGATCTGATCTGATCTGATCT 1057

```

Oy	1021	GATGCTGTCGAAACAACATGCGGTCAACAAGCTTGAGAAAGTATACGAAACAGAGCACTCG	1089
Db	1058	GATGCAAGTAGGTACTACTTGGCGGCAAGCTCTTGAAAGGTGTTGGAACTGAAACACTCT	1117
Oy	1081	GATATCTTTCGAGTACCCCTTCAGAAACAAAAAGGAAATGTTTCGAAAATGATCTCAGA	1140
Db	1118	CACATCTTACAGATTCCTCTTATAGACTGAAACGGAAATGTTTCGAAATGGATCTCAAGA	1177
Oy	1141	TTTGAAAAAGTCTGGCCATPACTTGGAAACCTACACAGAGATGTTGCTCATGAATCTCC	1200
Db	1178	TYCG---AAGCTGGCCCTACTCTGAAACTTACACCGAGGATGTTGCTCATGAGCTTGCA	1234
Oy	1201	AAAGGTTGCAACGGACGCCAGATGTGATTCATTCGGAAACMCAAGGACGGCAATATCGCTC	1266
Db	1235	AAAGGTTGCAAGGGCAAGCCAGATCTGATTCGTTGAAACCTACAGTGAAGAAACATTGTT	1294
Oy	1261	GCTCTCTTGTCTGCACATTAATTAAGTGTCAACAGTGCACACATCGCCCATGCTTTGGAG	1320
Db	1295	GCTCTTTGTTGGCACTAATAATTAAGTGTCACTGATGTACATTGTCATGCACTTGAG	1354
Oy	1331	AAAGCAAAATATCCAGATTCAGATTTCTATTGGAAGAACTTGAAGACAATPACATTTC	1380
Db	1355	AAAGCAAAATATCCGAAATCCGACATTTCGGAATAAACTGGAAGAGATPACACATTC	1414
Oy	1381	TCTTGCCAAATTAAGCGTGAATCTTTTGCAATGAACCAATACAGATTTCAATATCCAGT	1440
Db	1415	TCTTGCCAAATTAAGCGGATTAATTTGCAATGAACCAAGATTTCAATATPACAGT	1474
Oy	1441	ACTTTCAGAAATTCGAGAAAGCAAGACACTGTGTGTCATATAGAGGCACACTGCT	1500
Db	1475	ACATTTCAAAGAGATTGCTGGAAGCAAGACACTGTGTGCAAGTATGATCTCAACAGCC	1534
Oy	1501	TTCACTCTTCTCGGTCTCTACCGTGTGTACATGATTCATGTGTGATTCCAAATTC	1560
Db	1535	TYCACCTCTTCCGGACTCTACCGTGTGTGCAAGTATGATGTGCTTGATTCAAAATTC	1594
Oy	1561	AACATTGTTTCCCGGAGTCTGATATGAGAGATPACTTCCCTTACACCGAAGAGAGCGG	1620
Db	1585	AACATTGTTCTCTCCGGAGCTGATCAAAACATTTACTTCCCCACACCGAAACTAGCCGC	1654
Oy	1621	AGGTTGAAGCAATTTCCATCTGAGATCGAAGACCTTCTTTACACCAAGTTGAGATGAA	1680
Db	1655	AGGTTGACCTCTTTCACACCGAAATCGAAGAGCTTCTCTACAGTTCAGTGGAAGACAA	1714
Oy	1681	GAACTTATGTGTGCTCAATGACCGCAACAGCCAAATTCGTGTACATATGCCAAGCTT	1740
Db	1715	GAACTATATGTGTGCTGAGAGACCGCTCGAAGCCAAATTAATCTTCAACATGTGATGTTG	1774
Oy	1741	GATCGTGTCAAGAACTTAAACCGGACTCGTGCAGTGTGTCGGCAAGAACCCAAAGTTGCGT	1800
Db	1775	GACCTGTGAAAGACATCAACAGACTGTGTGAGTGTACGGTAAAGAACGCAAGCTGAGG	1834
Oy	1801	GAGTTGAGTAACTCTGTAGTGTGTAGTGTGTATGAGCCAAAGAACTTAAAGATTTGGAA	1866
Db	1835	GAGTTGTGGAACCTCGTGATCGTGTCTGTGAGACAGAGAAAGAGAGTCAAGAACTTGGAA	1894
Oy	1861	GAGAAAGCTGAATTAAGAAAAATGTTTGAAGCTGATCGCAAGTACAACTTGAACGGCCAA	1920
Db	1895	GAGAAAGCGGAGATGAAGAAAGTATAGCTGATCGAAGCGTACAACTTGAACGGGCAA	1954
Oy	1921	TTCAATGAGATTCATCTCAATGAAGAAATCCCAATGTGTAACCTTACCGATPACATT	1980
Db	1955	TTCAATGAGATTTCACTCAATGAAGCCGTGTGAGGAATGAGAGCTGTACCGTGTATC	2014
Oy	1981	TGCGACACGAAGGTGCTTTGTACAGCTGTGATTTGAGAGCTTTGATTTGACAGTT	2040
Db	2015	GCCGACACCAAGGGGTCTTTGTGTGACGCTGACGTACGAGGCAATTTGGTTGACAGTG	2074
Oy	2041	GTTGAGGCAATGACTTGGCGGTTTGGCAACTTCGCAACCTGTAAACGTGTGACAGCCGAG	2100
Db	2075	GTTGAGGCAATGACTTGGCGGTTTGGCAACTTTGCAACATGCAATGAGTGTGACCTGTGAG	2134

Oy		2101	ATTAATTGTCATGGGAAATCTGGTTTCAACATTGATCCTTACCATGGATGAATCAAGTGCT	2168
Db		2135	ATCATGTCCATGAAAATCTGGGTTTCCAATTGATCTTACCACGGTAGCCGCTGCT	2194
Oy		2161	GACATACTCGTCGATTTCTTTGAAAAAGTGTAGAAMAATGCATCTCACTGGGATTAAGATC	2220
Db		2195	GACCTCTGTGTAGTGTTCTTTTGAAAGAGTCAAGGTGACCCATCTCACTGGGACAAGATC	2254
Oy		2221	TCCCAAGAGAGCTTGAAACGAATPAGAGAGAAAGTATPACATGAAGATTTACTCGGAGAGA	2280
Db		2255	TCTCAACCCGGTCTCCCAACGATATTAAGAGAAAGTACACATGGCAAATTTACTCGCACAGG	2314
Oy		2281	CTATTGACCCGTGACAGAGTGTATGAGATTCGTGGAGACATGTTTCCAACTTTGAACGCCGT	2340
Db		2315	CTTCTACTCTTCACCGGTGTGTATGTTTTCTGGAAGCATGTCTTAACCTCGGATGCCGT	2374
Oy		2341	GAGAGTCGTGTTAACCCTTGAGATGTTTATGCTCTTAAGTACCGTAACTGCGCTGAATCA	2400
Db		2375	GAGGTGCGGGCTATCTTGAGATGTTCTATGCTCTCAAGTACCGCAAAATTTGGCTGAGTCT	2434
Oy		2401	GTTTCATTGGCAGAGAGTAAATTGAACCTGTTAAATAA	2439
Db		2435	GTGCCCCCTTGCTGTGTAGTAACTGATGAGATGAAGAA	2473
RESULT 4				
AF030231				
LOCUS		2842 bp	mRNA	linear PLN 21-JAN-1998
DEFINITION		Glycine max sucrose synthase (SS) mRNA, complete cds.		
ACCESSION		AF030231		
VERSION		AF030231.1	GI:2606080	
KEYWORDS				
ORGANISM		Glycine max (soybean)		
SOURCE		Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.		
REFERENCE		1 (bases 1 to 2842) Zhang,X.-Q., Verma,D.P.S., Patil,S., Arredondo-Peter,R., Miao,G.-H., Kuismannen,R., Klucas,R.V., and Chollet,R. Cloning of a full-length sucrose synthase cDNA from soybean (Glycine max) root nodules (Accession No. AF030231) (PCR 97-173) Plant Physiol. 115, 1729 (1997)		
JOURNAL		2 (bases 1 to 2842) Zhang,X.-Q., Verma,D.P.S., Patil,S., Arredondo-Peter,R., Miao,G.-H., Kuismannen,R., Klucas,R. and Chollet,R. Direct Submission Submitted (16-OCT-1997) Biochemistry, University of Nebraska-Lincoln, George W. Beadle Center, Lincoln, NE 68588, USA		
AUTHORS		Location/Qualifiers		
TITLE		1. .2842		
JOURNAL		/organism="Glycine max"		
FEATURES		/mol_type="mRNA"		
SOURCE		/db_xref="taxon:3847"		
		/feature_type="nitrogen-fixing root nodules"		
		1. .2842		
gene		/gene="SS"		
CDS		72..2489		
		/EC_number="2.4.1.13"		
		/note="nodulin-100"		
		/codon_start=1		
		/product="sucrose synthase"		
		/protein_id="AAC39323.1"		
		/db_xref="GI:2606081"		
		/translation="MATDIRTRVSHLRBRLDETITLNRNEITALSRLEAKGKIIQH HOYIAFEELPEENROKLTDGAGEYLRSQEAIVLPWVALAVRPFGWEYLRVNV HAUVEELOPAEYLHPRELVLDDSSNGNFYLEDPFPNAPRPTLNKSINGVOFL NRSLAKTLPHDKESLHPLEFRLIHLVKGTMLNDRIOMDPLAOHLBRAREYLGTV PERTPYSEBFHKOEIGLERGMKDNNRVLUESQLDLDEADPCLTEFLFRIRIPV FNVLILSPGYFAQDNVLNGFPGDGQVVYLLDVDRALNMENLHRIKOGGLIVPRILI ITFLPDVAVTTCGRLEKVGFGHREHSHTLVRPTEKGIVRKMIISREWVPYLETYE		

QY 1981 TCGCAGCAGAAAGTGCCTTGTGTAACAGCTGCATGTAATGAAGCTTTGATGACAGCTT 2040
DB 2049 TCGCAGCAGCAGAGGCTGCTTGTGACAGCTCTGTATACGAGGCTTTTGGTTGACAGTGT 2108
QY 2041 GTGAGGCAATGACTGTGGCTTGGCCAACTTCGCAACCTGTAAACGGGAGCCAGCCGAG 2100
DB 2109 GTTAAAGGCAATGACTGTGGCTTGGCCAACTTCGCAACCTGTAAACGGGAGCCGCTGAG 2168
QY 2101 ATTATTTGTCATGGGAAATCTGTTTCAACATTCCTTACATGATCAAGTCAAGCTCT 2160
DB 2169 ATCATTTGTCACGGCAAGCTGCGCTTCCACATTCCTTACATGATCAAGTCAAGCTCT 2228
QY 2161 GACATCTCTGTCATTTCTTTGAAAAGTGAAGAAAGATTCATCTCACTGGGATGAATC 2220
DB 2229 GATCTCTCTGTCATTTCTTTGAGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2288
QY 2221 TCCCAAGGAGCTTGAAGCAATGAGGAGAGTATACATGAGGAGATTTACTCGGAGAGA 2280
DB 2289 TCAAAAGGCTGCTTCCAGCGTATGAGAGAGTACATGAGGAGGAGGAGGAGGAGGAGG 2348
QY 2281 CTATTTGACCTGACAGAGAGTATGAGATTTCTGAAAGCATGTTTCCAACTTGAAGCGCT 2340
DB 2349 CTTCTCACTCTCAACCGGTGTATGAGCTTCTGAAAGCATGTTTCTAAGCTTGAAGCGCT 2408
QY 2341 GAGAGTGTCTGTAACCTTGAAGATTTTATGCTTAAAGTAAAGCTGAGCTGAAATCA 2400
DB 2409 GAGAGCGCGCTCTATCTGAGATGTTTATGCTCTCAAGTACCGCAATTTGCTGAGTCT 2468
QY 2401 GTTCCATTTGAGAGAGAGTAAATGA 2426
DB 2469 GTGCGCTTGTCTGCTGAGTAAATGA 2494

RESULT 5
AF315375 2738 bp mRNA linear PLN 02-DEC-2002
LOCUS AF315375
DEFINITION Phaseolus vulgaris sucrose synthase mRNA, complete cds.
ACCESSION AF315375.1 GI:25990359
VERSION
KEYWORDS
SOURCE
ORGANISM
Phaseolus vulgaris
Phaseolus vulgaris
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Phaseolus.
REFERENCE 1 (bases 1 to 2738)
Camas,A., Silvente,S. and Lara,M.
Expression of the gene for sucrose synthase during development of
common bean nodules
unpublished
JOURNAL 2 (bases 1 to 2738)
Camas,A., Silvente,S. and Lara,M.
Direct Submission
Submitted (20-OCT-2000) Programa de Biología Molecular de Plantas,
Centro de Investigación sobre Fijación de Nitrógeno, Av.
Universidad S/N, Cuernavaca, Mor. 62210, Mexico
location/Qualifiers
FEATURES
Source
1. 2738
/organism="Phaseolus vulgaris"
/mol_type="mRNA"
/db_xref="taxon:3885"
/clone="PVSS"
/tissue_type="nodules"
52. 2469
/BC_number="2.4.1.13"
/note="SS; D-fructose 2-glucosyltransferase; UDP glucose;
sucrose cleavage; predicted 92 kDa"
/codon_start=1
/product="sucrose synthase"
/protein_id="AA076498.1"
/db_xref="GI:25990360"
/translation="MAADRLTRVHSLRELDLTLNANREILALSLRIEAKGTGLIOLH

CDS
/organism="Phaseolus vulgaris"
/mol_type="mRNA"
/db_xref="taxon:3885"
/clone="PVSS"
/tissue_type="nodules"
52. 2469
/BC_number="2.4.1.13"
/note="SS; D-fructose 2-glucosyltransferase; UDP glucose;
sucrose cleavage; predicted 92 kDa"
/codon_start=1
/product="sucrose synthase"
/protein_id="AA076498.1"
/db_xref="GI:25990360"
/translation="MAADRLTRVHSLRELDLTLNANREILALSLRIEAKGTGLIOLH

ORIGIN

Query Match 60.1%; Score 1577.2; DB 8; Length 2738;
Best Local Similarity 78.3%; Pred. No. 0;
Matches 1904; Conservative 0; Mismatches 524; Indels 3; Gaps 1;
QY 1 ATGGCTGAGCGTCTCACTCGCTCCAGAGTCTCCGAGCGTTTGAATGAGACCTT 60
DB 52 ATGGCGCTGATGCTTGAACCGCTTCAACAGTCTCCGTAAGAGCTTGAATGAAACCTG 111
QY 61 CTGCTCAAGAAACGAGATTTGGCTTCTCTCAAGATCGAGGCAAGAAAGAAAGAA 120
DB 112 TCTGCCAAACGAAACGAAATTTGGCCCTTCTGCAAGATAGAGCCAAAGGCGACCGGG 171
QY 121 ATTCTGCAACACCATCAATATTCTAGATTTGAAGCTATCCCTGAAAGAAACAGAAAG 180
DB 172 ATCTTGCAACCAATCAAGTCAATTTGCTGAGTGAAGAAATCCCGAGAGAGCAAG 231
QY 181 AAGCTCGATAGATGATTTTGAAGTGAAGCTAGTGAAGAGCATGCTT 240
DB 232 AAGCTATATGATGATGCTTGAAGAAATTTGAAGCTATCAAGAAAGCATGTTT 291
QY 241 CTTCAATGAGTGAAGTCTGCTGCTTCAAGGCTGCTGTTTGAAGTCAATTAAGT 300
DB 292 CCAACATGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 351
QY 301 AATGTCAGCGCTTGTGTTGAGAACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 352 AATGTCAGCGCTTGTGTTGAGAACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 411
QY 361 GAGCTGTTGATGAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 420
DB 412 GAATGTTGATGAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 471
QY 421 AACTCATCATTTCCCGCCCACTTTTCAAAATTCATGCTGTAAGTGTGAGTCTCTA 480
DB 472 ACGGATCTCTTCCCGCCCACTTTTCAAAATTCATGCTGTAAGTGTGAGTCTCTC 531
QY 481 AATGTCAGCGCTTGTGTTGAGAACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 532 AATGTCAGCGCTTGTGTTGAGAACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 591
QY 541 TTCTCAGATTCATTTGTCAAAGGCGCAAGATGTTGTAATGACAGAAATTCAGAAC 600
DB 592 TTCTCAGATTCATTTGTCAAAGGCGCAAGATGTTGTAATGACAGAAATTCAGAAC 651
QY 601 TTGATGCTTTTCAACATGTTTGAAGAAACGAGAGATGTTGTAATGACAGAAATTCAGAAC 660
DB 652 CTTGATGCTTTTCAACATGTTTGAAGAAACGAGAGATGTTGTAATGACAGAAATTCAGAAC 711
QY 661 GAGACACATGTCGCAATTCGAACCGGTTCAAGAAATGTTGTAATGACAGAAATTCAGAAC 720
DB 712 GAAACCCCTTCACTGAGATTTGACACAGATTCAGAAATTTGTTGAGAGAGGCTG 771
QY 721 GGTGACACGCAAGCGTGTGAGATGATCAATCTCTTTTGAATCTTTTGAAGCA 780
DB 772 GGTGACACGCAAGCGTGTGAGATGATCAATCTCTTTTGAATCTTTTGAAGCA 831
QY 781 ACTGATCTTTGACACCTTGAAGATTTCTTTGAGAAATTCCTCATGCTGTTCAATGTTG 840


```
gene
1. .2728
/gene="bucsl"
43. .2460
/gene="bucsl"
/EC_number="2.4.1.13"
/codon_start=1
/product="sucrose synthase"
/protein_id="CAB40794.1"
/db_xref="gi:4584690"
/db_xref="GOA:O9T0M6"
/db_xref="SPTREMBL:O9T0M6"
/translation="MATERLTVHSLSKERLDETLTANRNELLALSRLKAGKGLIHH
HOVIAEPEIPEBDSROKLTGAPGEVLRSTOEALVLEPPVALAPRPPGIREYRVAV
NALVENIQLPAEFLKPEKELVDGSGANFVLEDPPEPTLSPPTLANKSIGNGOPL
NRHLASLFDKESLHPLLEFLRLHSYKTLMLNDR1IONDSLOHVLARKAEYLSLT
DPEPYSEFERFOEIGLERGMDTAERLEVS10LLDLLEAPDCTLEFLDR1PMV
FNVLISPHGVFADDDVLYGPDGGOVYVILDOVRALRESEMLSR1KKOGLD1IPRL1I
ITRLPRAVETTCORLEKYGTGHECHILRPVPRDTGIYARKW1SRBEVMPYETTYE
DVAEHLAELOQKFDLIVGNTSDNIVASLHAKLGVYQCTTIAHLEKTYKPSDITYM
KKEBEKHFSQCFADLEFANMHTDPIITSTOE1AGSDKVGQYSEHSTAFLPELYEV
VHGIDVDPKENIVSPGADQTIYEPYETSRRLTSFYPELEL1YSSEVNEEH1CVLK
DRNKPI1ETMARLDVRKNI1GLVEMYGKNA1LRLVNLVVAAGDRKESKOLEE1AM
KKWYGL1ETTYLNGOFPM1ISSOMNRVNRGEL1YRVIDPKGAFVGPAYEARGLTVEA
MATLPTPATLNGSPAR1IYHKGSPH1DPHGDRAAD1LYEPEK1KVDPSHDK1S
OGGLQRIEKTMTT1SORLL1LTVGVGFWMHVSNDRLRESRYL1EMFYALKRKLAE
SVPLAVE"
```

ORIGIN

Query Match 59.4%; Score 1559; DB 8; Length 2728;
Best Local Similarity 78.0%; Pred. No. 0; Mismatches 531; Indels 3; Gaps 1;
Matches 1890; Conservative 0;

```
QY 1 ATGGCTGAGCGTGTCTCTCATCGCGCTCCAGAGTCTCCGTGAGCGTTTGATGAGCCCTT 60
D 1 ATGGCTGAGCGTGTCTCTCATCGCGCTCCAGAGTCTCCGTGAGCGTTTGATGAGCCCTT 60
QY 43 ATGGCTACCGAAGCTTTGACTGTGTTTATAGCCCTCAAGAGAGGCTTGATGAACCTTA 102
D 43 ATGGCTACCGAAGCTTTGACTGTGTTTATAGCCCTCAAGAGAGGCTTGATGAACCTTA 102
QY 61 CTGGCTCAGAGAACGAGATTTTGGCCCTTGCTCTCAGAGATCGAGGCAAGAAAAGA 120
D 61 CTGGCTCAGAGAACGAGATTTTGGCCCTTGCTCTCAGAGATCGAGGCAAGAAAAGA 120
QY 103 ACTCTATAGAGATGAATTTTGGCCCTTCTTTCAGGCTTGAAAGCAAGGAAAAGGA 162
D 103 ACTCTATAGAGATGAATTTTGGCCCTTCTTTCAGGCTTGAAAGCAAGGAAAAGGA 162
QY 121 ATTCTGCAAGCCATCAATTAATCTAGATTGAAGTATCCCTGAAAGAAACAGAAAG 180
D 121 ATTCTGCAAGCCATCAATTAATCTAGATTGAAGTATCCCTGAAAGAAACAGAAAG 180
QY 163 ATTTTGCAACCAACATGATATGCTGAGTTTGAGAAATTCCTGAAGATGATGACAG 222
D 163 ATTTTGCAACCAACATGATATGCTGAGTTTGAGAAATTCCTGAAGATGATGACAG 222
QY 181 AAGCTCGTAATGTGATTTTGTGAAGTATGAAAGCTAGTCAAGAACGATCGTTTG 240
D 181 AAGCTCGTAATGTGATTTTGTGAAGTATGAAAGCTAGTCAAGAACGATCGTTTG 240
QY 223 AAGTTGACTAGTGTGATTTGGTGAATTTTGAAGTCAACAGAGAAAGCAATGTTTG 282
D 223 AAGTTGACTAGTGTGATTTGGTGAATTTTGAAGTCAACAGAGAAAGCAATGTTTG 282
QY 241 CCTCCATGAGTTGCACTTGCTGTTCGTCAGAGCGCTGTGTTTGGAGTACATTAAGTG 300
D 241 CCTCCATGAGTTGCACTTGCTGTTCGTCAGAGCGCTGTGTTTGGAGTACATTAAGTG 300
QY 283 CCACCATGAGTTGCACTTGCTGTTCGTCAGAGCGCAAGGTATTTGGAGTATCTGAGATA 342
D 283 CCACCATGAGTTGCACTTGCTGTTCGTCAGAGCGCAAGGTATTTGGAGTATCTGAGATA 342
QY 301 AATGTTCACGCGCTTGTTGTGAGGAATCACTGTGTGAGTATCTCCATTTGAAGAA 360
D 301 AATGTTCACGCGCTTGTTGTGAGGAATCACTGTGTGAGTATCTCCATTTGAAGAA 360
QY 343 AATGTGACGTCTGTGTGTGAATAATTGCAACCTGTGAGTCTCAAAATTCAGAGAA 402
D 343 AATGTGACGTCTGTGTGTGAATAATTGCAACCTGTGAGTCTCAAAATTCAGAGAA 402
QY 361 GAGCTTGTGAGAAAGTTCAAAATGAAACTTTGTTTGAATGAAATTTGAGACCCCTTC 420
D 361 GAGCTTGTGAGAAAGTTCAAAATGAAACTTTGTTTGAATGAAATTTGAGACCCCTTC 420
QY 403 GAACCTGTGATGAGAAAGTCTAATGGAACCTTTGTGCTGAGTTGSACTTTGAACATTT 462
D 403 GAACCTGTGATGAGAAAGTCTAATGGAACCTTTGTGCTGAGTTGSACTTTGAACATTT 462
QY 421 AACTCATATTCCTCCCGCCCACTCTTCAAAATTCATGTTAAATGTTGGAAGTTCTTA 480
D 421 AACTCATATTCCTCCCGCCCACTCTTCAAAATTCATGTTAAATGTTGGAAGTTCTTA 480
QY 463 ACTCATATTCCTCCCGCTCTAATCTCAACAGTCAATTTGGAATTTGGTGTCAATTCCTT 522
D 463 ACTCATATTCCTCCCGCTCTAATCTCAACAGTCAATTTGGAATTTGGTGTCAATTCCTT 522
QY 481 AATGTACCTTTGGGGAATAATGTTCCATGCAAGAGAGACAGCAACCTTTGCTGAA 540
D 481 AATGTACCTTTGGGGAATAATGTTCCATGCAAGAGAGAGACAGCAACCTTTGCTGAA 540
QY 523 AATGCAACCTTTTGTCTTAACTCTTCATGCAAGAGAGATTACATCCACTTTGAA 582
D 523 AATGCAACCTTTTGTCTTAACTCTTCATGCAAGAGAGATTACATCCACTTTGAA 582
QY 541 TTCTCTAGAGTCAATTTGTCAAGAGGCAAGAACTGATGTTGAATGACAGAAATTCAGAAC 600
D 541 TTCTCTAGAGTCAATTTGTCAAGAGGCAAGAACTGATGTTGAATGACAGAAATTCAGAAC 600
QY 583 TTTCTCAGACTTCAAGCTCAAGGGAAGACATTTGATGTTGATGACAGAAATTCAGAAAC 642
D 583 TTTCTCAGACTTCAAGCTCAAGGGAAGACATTTGATGTTGATGACAGAAATTCAGAAAC 642
QY 601 TTGAATGCTTTCACATGTTTGTGAGAAAGACAGAGATCTTGATCCCTTACTCTCT 660
D 601 TTGAATGCTTTCACATGTTTGTGAGAAAGACAGAGATCTTGATCCCTTACTCTCTCT 660
```

```
D 643 CTGATTTCTCTTCAACATGTTCTGAGGAAGCTGAAAGATTTCAACACAAATGATCTT 702
QY 661 GAGACACCATGTCGCGCAATTTGAAACACCGGTTCCAGAAAATCGGTTTGAAGAAGTTGG 720
D 661 GAGACACCATGTCGCGCAATTTGAAACACCGGTTCCAGAAAATCGGTTTGAAGAAGTTGG 720
QY 703 GAAACACCATATCTCAGAAATTTGAAACAGAGTTCCAGAGATGGTTTGGAGAGAGTTGG 762
D 703 GAAACACCATATCTCAGAAATTTGAAACAGAGTTCCAGAGATGGTTTGGAGAGAGTTGG 762
QY 721 GGTGACACCGGAGAACGCGTGTCTGAGATGATTCACATCTCTTTGAAATCTTCTTGAGCA 780
D 721 GGTGACACCGGAGAACGCGTGTCTGAGATGATTCACATCTCTTTGAAATCTTCTTGAGCA 780
QY 763 GGAGACACCCGAGACCGCGTCTCGAGTCCATCCAACTTCTCTTGATCTTCTGAGGCT 822
D 763 GGAGACACCCGAGACCGCGTCTCGAGTCCATCCAACTTCTCTTGATCTTCTGAGGCT 822
QY 781 ACTGATCTTGACACCTTTGAGAAAGTTCCCTGGAGAAATCCCAATGAGTTCATAGTTGTG 840
D 781 ACTGATCTTGACACCTTTGAGAAAGTTCCCTGGAGAAATCCCAATGAGTTCATAGTTGTG 840
QY 823 CCCGACCTTTGACACCTTTGAGACTTCTCTGATGAATATCCCAATGAGTTCATAGTTGTG 882
D 823 CCCGACCTTTGACACCTTTGAGACTTCTCTGATGAATATCCCAATGAGTTCATAGTTGTG 882
QY 841 ATTCTCACTCCCAACGAGATCTTCTGCTCAAGACAAATGTTTGGGATATCCGACACCGGT 900
D 841 ATTCTCACTCCCAACGAGATCTTCTGCTCAAGACAAATGTTTGGGATATCCGACACCGGT 900
QY 883 ATCTTTCTCTCTCATAGTTTACTTTGCTCAAGATGATGTCTTGGATATCCCTGATACGTGGA 942
D 883 ATCTTTCTCTCTCATAGTTTACTTTGCTCAAGATGATGTCTTGGATATCCCTGATACGTGGA 942
QY 901 GGGCAGGTTGTTTATCATCTTTGATCAAGTCCGAGCTTTGAGAAATGAGATGCTTCTCGT 960
D 901 GGGCAGGTTGTTTATCATCTTTGATCAAGTCCGAGCTTTGAGAAATGAGATGCTTCTCGT 960
QY 943 GGCCAGGTTGTTTATCATCTTTGATCAAGTCCGAGCTTTGAGAAATGAGATGCTTCTCGT 1002
D 943 GGCCAGGTTGTTTATCATCTTTGATCAAGTCCGAGCTTTGAGAAATGAGATGCTTCTCGT 1002
QY 961 ATAAAGCAAGAGACTCAACATCAACCCCTCGAATCTCTCATTTATTAAGATCTTCTCT 1020
D 961 ATAAAGCAAGAGACTCAACATCAACCCCTCGAATCTCTCATTTATTAAGATCTTCTCT 1020
QY 1003 ATTAAGAAACAGGCTTGATATCATCTCTCGCATTTCTCATTAACACCCGCTTCTCTCC 1062
D 1003 ATTAAGAAACAGGCTTGATATCATCTCTCGCATTTCTCATTAACACCCGCTTCTCTCC 1062
QY 1021 GATGCTGTGGAACAACATGCGGTCAACGACTTGAGAAAGTATPCGGAACAGAGCACTCG 1080
D 1021 GATGCTGTGGAACAACATGCGGTCAACGACTTGAGAAAGTATPCGGAACAGAGCACTCG 1080
QY 1063 GATGCAATCGGAACGACTTGTGGCCACGACTTGAGAAAGTATPCGGAACAGAGCACTCG 1122
D 1063 GATGCAATCGGAACGACTTGTGGCCACGACTTGTGGAAAGTATPCGGAACAGAGCACTCG 1122
QY 1081 GATATTTCTGAGTACCTTTCAAGAACGAAAGGAAATTTTGGAAATGAGATCTCAAGA 1140
D 1081 GATATTTCTGAGTACCTTTCAAGAACGAAAGGAAATTTTGGAAATGAGATCTCAAGA 1140
QY 1123 CACATTTCTGAGTACCTTTCAAGATACGAAAGGAAATTTTGGCAAGTGGATCTCAGA 1182
D 1123 CACATTTCTGAGTACCTTTCAAGATACGAAAGGAAATTTTGGCAAGTGGATCTCAGA 1182
QY 1141 TTTGAAAAATCTGCGCATCTTGAGAAACCTTACACAGAGATGTTGCTCATGAATCTCC 1200
D 1141 TTTGAAAAATCTGCGCATCTTGAGAAACCTTACACAGAGATGTTGCTCATGAATCTCC 1200
QY 1183 TTTG--AAGTCTGCGCATCTTGAAGAACTTGAACCTGAGATGTTGCTCATGAGCTTGC 1239
D 1183 TTTG--AAGTCTGCGCATCTTGAAGAACTTGAACCTGAGATGTTGCTCATGAGCTTGC 1239
QY 1201 AAAGAGTTGACAGGACGCGAGATCTGATCATCGAAACNAACGACGCAATATGCTC 1260
D 1201 AAAGAGTTGACAGGACGCGAGATCTGATCATCGAAACNAACGACGCAATATGCTC 1260
QY 1240 AAAGAGTTGACAGGACGCGAGATCTGATCATCGAAACNAACGAGTATGAAATGTT 1299
D 1240 AAAGAGTTGACAGGACGCGAGATCTGATCATCGAAACNAACGAGTATGAAATGTT 1299
QY 1261 GCTCTCTTGTCTCGACATTAATTAAGTGTGACACAGTGCACATGCCATGCTTTGGAG 1320
D 1261 GCTCTCTTGTCTCGACATTAATTAAGTGTGACACAGTGCACATGCCATGCTTTGGAG 1320
QY 1300 GCTCTTGTGTGACATTAATTAAGTGTGACATGAGTGTGATGCTGATGCACTGGAG 1359
D 1300 GCTCTTGTGTGACATTAATTAAGTGTGACATGAGTGTGATGCTGATGCACTGGAG 1359
QY 1321 AAGACAAAATATCCAGATTCAGATATCTATTTGAGAAAGCTTGAAGCAATATCATTTTC 1380
D 1321 AAGACAAAATATCCAGATTCAGATATCTATTTGAGAAAGCTTGAAGCAATATCATTTTC 1380
QY 1360 AAGACTAAGTACCCGCAATCCGATTTACTGGAAGAAATTCGAAGAGAGATCACCTTC 1419
D 1360 AAGACTAAGTACCCGCAATCCGATTTACTGGAAGAAATTCGAAGAGAGATCACCTTC 1419
QY 1381 TCTTGCAATTTACAGTGTCTTTTGTGCAATGAAACATACAGATTTTATCATATCACAGT 1440
D 1381 TCTTGCAATTTACAGTGTCTTTTGTGCAATGAAACATACAGATTTTATCATATCACAGT 1440
QY 1420 TCTGCGCAATTTACGCTGATCTTTTGTGCAATGAAACACAGATTTTATCATATCACAGT 1479
D 1420 TCTGCGCAATTTACGCTGATCTTTTGTGCAATGAAACACAGATTTTATCATATCACAGT 1479
QY 1441 ACTTTCAGAGAAATTTGAGAGAGCAAGCACTGTTGTCAATACGAGAGCAACATGCT 1500
D 1441 ACTTTCAGAGAAATTTGAGAGAGCAAGCACTGTTGTCAATACGAGAGCAACATGCT 1500
QY 1480 ACTTTCAGAGAGATTTGAGAGAGCAAGGTTGAGAGATGAGAGTCAACATGCTC 1539
D 1480 ACTTTCAGAGAGATTTGAGAGAGCAAGGTTGAGAGATGAGAGTCAACATGCTC 1539
QY 1501 TTGACTTCTCTGCTCTCTACCGTGTGTACAGATATGATGATGTTGATATCCAAATTC 1560
D 1501 TTGACTTCTCTGCTCTCTACCGTGTGTACAGATATGATGATGTTGATATCCAAATTC 1560
QY 1540 TTTACTCTTTCAGAGACTTACCGGTGTGTGACAGGATATGATGCTTTGATATCCAAATTC 1599
D 1540 TTTACTCTTTCAGAGACTTACCGGTGTGTGACAGGATATGATGCTTTGATATCCAAATTC 1599
QY 1561 AACATGTTTCCCTGAGTGTGATATGAGATATATCTTCCCTTACACCGAAGAGAACCGG 1620
D 1561 AACATGTTTCCCTGAGTGTGATATGAGATATATCTTCCCTTACACCGAAGAGAACCGG 1620
QY 1600 AACATGTTATCTCCAGAGAGTGTATCAGACATTTATCTTCCCTTACACCGAAGAGAAC 1659
D 1600 AACATGTTATCTCCAGAGAGTGTATCAGACATTTATCTTCCCTTACACCGAAGAGAAC 1659
QY 1621 AGGTGAAGATTTTCAATCTGAGATGAGAGACCTTCTTACACCAAAGTTGAGATGAA 1680
D 1621 AGGTGAAGATTTTCAATCTGAGATGAGAGACCTTCTTACACCAAAGTTGAGATGAA 1680
QY 1660 CGATGATCATCTTCTTACCTGAAATGAGAGAGCTTCTTACAGCTCAGTTGAGAAATGAA 1719
D 1660 CGATGATCATCTTCTTACCTGAAATGAGAGAGCTTCTTACAGCTCAGTTGAGAAATGAA 1719
QY 1681 GAACATTAATGTGTGCTCATATGACCGGACAAAGCAATTTCTGTTCAAAATGCCAGCTT 1740
D 1681 GAACATTAATGTGTGCTCATATGACCGGACAAAGCAATTTCTGTTCAAAATGCCAGCTT 1740
```


D 505 AACGTCACCTTTCTGCCAACTCTTCATGACAGAGAGTTTGCATCCATTGGA 564
Q 541 TTCTCAGAGTCATTTGTACAGAGGCAAGAAATGATGTTGAATGACAGATTCAGAAC 600
D 565 TTCTCAGACTTCAAGCTACAGAGGGAAGCATTTGATGTTGAATGACAGATTCAGAAC 624
Q 601 TTGAATGCTCTTCAACATGTTTGAAGGAAGCAGAGGATCTTTGGTACCTTACCTCT 660
D 625 CCTGATCTCTTCAACATGTTTGAAGGAAGGCTGAGAGGATCTTGAAGGACGTTGCTCG 684
Q 661 GAGACACCATGTGCGAAATTCGAACAACGGTTCCGAAATCGGTTTGAAGAAGGTTGG 720
D 685 GATACACCGTACTCGAATTCGAACACAGGTTCCAGAGATTTGGTTGAGAGAGGTTGG 744
Q 721 GGTGACACCGGAGAACCGGCTCGAGATGATCCACTCTTTTGATCTTTGAGGCA 780
D 745 GGAGACACCGGAGACCGGCTCGAGTCCATTGAGCTTCTTATGATCTTTGAGGCT 804
Q 781 ACTGATCTTGCACCTTGAAGAGTTCTTTGGAGAAATCCCATGCTGTTCAATGTTGTG 840
D 805 CCTGATCTTGCACCTTGAAGAGTTCTTTGACAGAAATCCCATGCTGTTTAAATGTTGTC 864
Q 841 ATTGCACTCCCAACGATCTGCTGACAGACAAATGTTTGGGGTATCCGACACCGGT 900
D 865 ATCTCTTCTCTCAGGTTACTTTGCTCAAGATGATGTTTGGGATACCTTGATCCGGT 924
Q 901 GGGCAGGTTGTTTACATCTTGAATCAATCCGAGCTTGAAGATGAGATCTCTCGT 960
D 925 GGTGAGGTTGTTTACATTTGGATCAAGTTGCTGCTTGAAGAGAGATCTCAATGCG 984
Q 961 ATAAAGCAACAGACTCAATCAACCTCGAATCTCTATTTATCTAGCTTCTTCT 1020
D 985 ATTAAGAAACAGAGCTGATGATGTTCTCGCAATCTCATTAACACCGCTTCTCCCA 1044
Q 1021 GATGCTGCGGAAACAATGCGGTCACAGCTTGAAGAAATGATACGAAACAGACACTCG 1080
D 1045 GACGAGTCGGAACCTACTTGTGGCCACAGCTCGAAGAGTTTATGAAACCGAGATTCG 1104
Q 1081 GATATCTTTCGAGTACCTTTCAGAACAGAAAGGAAATGTTGCAAAATGATCTCAGA 1140
D 1105 CACATTTCTTGAGTTCCCTTCAGAGATCAGAAAGGAAATGTTCCCAAGTGAATCCGGT 1164
Q 1141 TTTGAAAAAGTCTGGCCATCTTGAACCTACACAGAGATGTTTGTCTCATGAATCTCC 1200
D 1165 TTCC---AAGTCTGGCCATCTAGAAACCTACACCGAGATGTTGTCTCATGACCTTGC 1221
Q 1201 AAAGAGTTGACGCGACCGCATCTGATCAACGAAACNAACGACGCGCATATGCTC 1260
D 1222 AAAGAGTTGACGCGACCGCATCTGATCAACGAAACNAACGATGAGAAATGTT 1281
Q 1261 GCTCTCTTGTGCAATAAATTAAGTGTCAACAGTGCACATCGCCCATGCTTTGGAG 1320
D 1282 GCTCTTGTGTTGGACATTAATTAAGTGTCACTAGTGTACTAATGTGCATGACCTTGG 1341
Q 1321 AAGCAAAATATCCAGATTCAGATTAATCTATTTGAAGAAAGCTTGAAGACAAATACATTC 1380
D 1342 AAGCAATAGTATCTGATCTGATATTAATCTGAAAAAATTCGAAGAAATATCACCTTC 1401
Q 1381 TCTTGCAAAATTAAGCTGATCTTTTGAAGAAACCATACAGATTCATCATGACCGT 1440
D 1402 TCTTGCAAAATTAAGCTGATCTTTTGAAGAAACCATACAGATTCATCATGACCGT 1461
Q 1441 ACTTTCAGAAATATTCAGAGACAGACAGCTGTTGTCAATACAGAGCAGACCTGCT 1500
D 1462 ACTTTCAGAGATTTGCTGGAAGAGATCTGTTGAAGAGTATGAGAGTCAACCTGCC 1521
Q 1501 TTCACTCTTCTGCTCTTACCGTGTGTATAGTGTATCATGATGTTGATCCCAATTC 1560
D 1522 TTCACTCTTCTGCGGACGTGTACCGTGTGTATGATGATGATGATGATGATGATGATG 1581
Q 1561 AAGATGTTTCCCTGCTGATATGAGATATATCTCCCTTCAACCGAGAGAGAGCGG 1620
D 1582 AAGATGATCTCTGAGAGCTGATGACCAATTAATCTCCCTTCAACCGAGAGAGAGCGG 1641

Q 1621 AGTTGAAGATTTCCATCTGATGAGAGACCTTTTACACCAAACTTGAGATGCA 1680
D 1642 AGTTGAAGATTTCCATCTGATGAGAGACCTTTTACACCAAACTTGAGATGCA 1701
Q 1681 GAACTATATGTTGCTCAATGACCGGCAAGGCAATTCGTTTCAATGCAAGGCTT 1740
D 1702 GAACATATATGTTGCTCAATGACCGGCAAGGCAATTCGTTTCAATGCAAGGCTT 1761
Q 1741 GATGTTGCAAGAACTTAAACCGGACTGCTGAGTGTGCGGCAAGAAACCAAGTTGCT 1800
D 1762 GACGTTGCAAGAACTTAAACCGGACTGCTGAGTGTGCGGCAAGAAACCAAGTTGCT 1821
Q 1801 GAGTTGCTTAACTCTGATGTTGATGAGTGTGATGAGGCAAGAACTTAAAGATTTGGA 1860
D 1822 GAGTTGCTTAACTCTGATGTTGATGAGTGTGATGAGGCAAGAACTTAAAGATTTGGA 1881
Q 1861 GAGAGGCTGAAATGAAAGAAATGTTTGAAGTGTGAGCAAGTCACTTGAACGCGCAA 1920
D 1882 GAGAGGCTGAAATGAAAGAAATGTTTGAAGTGTGAGCAAGTCACTTGAACGCGCAA 1941
Q 1921 TTCAATGATATATCATCTCAATGAAACAGATCCGAATGTTGAATCTTACCGTATCAT 1980
D 1942 TTCAATGATATATCATCTCAATGAAACAGATCCGAATGTTGAATCTTACCGTATCAT 2001
Q 1981 TCGCACAGAAAGGCTCTTGTATACAGCTGATGATGATGATGATGATGATGATGATG 2040
D 2002 TGTGACACAAAGAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATG 2061
Q 2041 GTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
D 2062 GTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2121
Q 2101 ATTAATGCTGAGGAAATCTGTTTCAATGATGATGATGATGATGATGATGATGATG 2160
D 2122 ATTAATGCTGAGGAAATCTGTTTCAATGATGATGATGATGATGATGATGATGATG 2181
Q 2161 GACATCTGCTGATTTCTTTGAAAGTGTAAAGATGATGATGATGATGATGATGATG 2220
D 2182 GATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2241
Q 2221 TCCCAAGAGGCTTGAAGAAATGAGGAGAGATGATGATGATGATGATGATGATGATG 2280
D 2242 TCCCAAGAGGCTTGAAGAAATGAGGAGAGATGATGATGATGATGATGATGATGATG 2301
Q 2281 CTATGACCTGACAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATG 2340
D 2302 CTATGACCTGACAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATG 2361
Q 2341 GAGAGTGTGCTTACCTTGAAGATGTTTATGCTTAAAGTGTGATGATGATGATGATG 2400
D 2362 GAGAGTGTGCTTACCTTGAAGATGTTTATGCTTAAAGTGTGATGATGATGATGATG 2421
Q 2401 GTTCCATTTGACAGAGATG--AATGAACTGTAAATTAACATTTGGCGGCTTTTCTT 2458
D 2422 GTTCCATTTGACAGAGATG--AATGAACTGTAAATTAACATTTGGCGGCTTTTCTT 2481
Q 2459 GGAATATATATCTGTTTGTATTAATTAATGAGAGCTCTTGTATTTCAATCTGT 2518
D 2482 ATCTGAGTGTGTTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2541
Q 2519 CTATTTCTTTTCTTTTGTGCGGCAATGTTTGAACATGAGGTTGTGCGGCTTCAAT 2578
D 2542 TGTTGATTAAGCTTTGATTAAGAAATGCAATGCTTTTCTTTTGTGATGATGATA 2601
Q 2579 CCAATTAATATGATGATTTGTTTCAAAAAAATTAATTAATTAATTAATTAATTAAT 2625
D 2602 TGTTGATTAAGCTTTGATTAAGAAATGCAATGCTTTTCTTTTGTGATGATGATA 2648

RESULT 8
AF049487
LOCUS AF049487 2760 bp mRNA linear PLN 01-JUN-1998

Db 259 AAACGATGATGGGAGCTTGGGAGAGTCCGAGAGCCACTCAGGAATCCATAGTTTG 318
Qy 241 CCTCAGAGGATGACATCTGCTGCTGCTCAGAGGCTGCTGTTGGAGATCAATAGATG 300
Db 319 CCCCCATGGGTTGCTCTGCTGCTGCTCAGAGGCGAGGCTGCTGGAATATATTAATCTC 378
Qy 301 AATGTCACGCGCTTGTGTTGAGGAATCACTGTTGCTGATATCTCAGCTTCAAGGA 360
Db 379 AATGTCATGACATGTTGTTGAGAGCTTCAAGTTGCTGATATTTGCACTTGAAGAA 438
Qy 361 GAGCTTGTGAGAGATTCATATGAAATCTTGTGTTGAAATGATTTGAGCCCTTC 420
Db 439 GAATCTGAGATGGAAGTTGCAAGGTAACTTGCTTGAAGCTTGAATTTGAGCAATTT 498
Qy 421 AACTCATCATTCCTCCCGCCCACTCTTCAAAATCATTGATATGCTGAGATTCCTA 480
Db 499 AATGCTTCCTTCTCGGCTTACTCTCTCAAGTCATTTGGCAATGGCGTTGAGTTCTCTC 558
Qy 481 AATGTCACCTTTGGGCAAAATGTTCCATGACAAAGGAGACATGACCTTGTCTGAA 540
Db 559 AACGTCACCTCTCGCAAGACTATTCATGATTAAGAGATATGCAAGCCATGCTGAC 618
Qy 541 TTCTCAGAGTCATTTGTCAGAAAGGCAAGACATGATGTTGAAATGACAGAAATCAGAAC 600
Db 619 TTCTCAGAGTCATCTGCTACAGGGAAGACATGATGTTGAATGATGAAATTCAGAAC 678
Qy 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGACAGATATCTTGGATCCCTACCTCT 660
Db 679 CTGATGCTCTTCCAGCATGTGCTGAGAGGACAGAGATATCTAGCAACCTTGACCT 738
Qy 661 GAGACACATGTCGCAATTTGAAACAAGGTTTCCAGGAATCGGTTTGAAGAAAGTTGG 720
Db 739 GAAACACCTTACTCTGAGATTTGCAACAGTTCCAGGAGATTTGGCTTGAAGAGAGATGG 798
Qy 721 GGTGACACCGCAGAACCGCTGCTCAGATGATCCAACTCTTGTGATCTTCTTGAAGCA 780
Db 799 GGAACACCTGCGAGCGGTCTGGAATGATCCAACTCTTGTGATCTCTGAGAGCC 858
Qy 781 ACTGATCTTGCACCTCTTGAAGAGTTCTTGGAGAAATCCCAATGCTGTTCAATGTTGTG 840
Db 859 CCGATCTCTGCAATTCGAGAAAGTCTTGTGTAATCCCAATGCTTCTTAAATGTTGTCTC 918
Qy 841 ATTCTCATCTCCCAAGGATCTTGGCTCAAGACATGTTTGGGGTATCCGCAACCGGT 900
Db 919 ATCTTTGCCCCAGGTTACTTTCGACAAAGATATGCTTGGGTTACCTCGATATCTGCT 978
Qy 901 GGGCAGGTTGTTTACATCTTGAATCAAGTCCGAGCTTGGAGAAATGAGATGCTCTCCGT 960
Db 979 GGGCAGGTTGTTTACATCTTGAATCAAGTCCGAGCTTGGAGAAATGAGATGCTCTCCGT 1038
Qy 961 ATAAAGCAACAGAGCTCAACATCAACCTCGAATCTCTATTAATTAAGATCTTCTCT 1020
Db 1039 ATAAAGCAACAGAGCTCAACATCAACCTCGAATCTCTATTAATTAAGATCTTCTCTCA 1098
Qy 1021 GATGCTGTCGGAACAAACATGCGGTCAACGACTTGAAGAAATGATACGGAACAGAGACTCG 1080
Db 1099 GATCAGAGGAAACCACTGCAACCGAGCTCTGAGAAAGTATTTGGAGACAGAGACTCC 1158
Qy 1081 GATATTTCTTCAGATCCCTTCAAGACAGAAAGGAAATTTGTCAGAAATGATCTCAGA 1140
Db 1159 CACATTTCTGAGTTTCAATTCAGAAAGGAAAGAAATGATCCGAAATGATCTCAGAA 1218
Qy 1141 TTTGAAAAAGTCTGCGCATACTTGGAAACCTTACACAGAGATGTTGCTCATGAATCTCC 1200
Db 1219 TTTG---AAGTCTGCGCTTACTTGAAGACTTACACCGAGATGTTTGCACAGAACTTACC 1275
Qy 1201 AAAGATTTGACGCGACGCGAGATCTGATCATGGAACACAGGACGGAATATGCTC 1260
Db 1276 AAAGATTTGACGCGAAACCGAGATCTGATCATGGAACACAGGACGGAATATGCTC 1335
Qy 1261 GCCTCTTGTGTCACATAAATTAAGTGTCAACAGTGCACATGCGCCATGCTTTGGAG 1320

Db 1336 GCCTCATTTGTCGACATAACTGGAGATTACAGATGTACTATGCTCATGCTTGGAG 1395
Qy 1321 AAGCAAAATATCCAGATTGATATCTATTTGAAAGAACTTGAAGCAATATACATTTTC 1380
Db 1396 AAAATTAATATCTGATTCAGATATTTATTTGAAAGATTTGATGACAGATATCACTTC 1455
Qy 1381 TCTTGCAGAAATTAACAGTATCTTTTTCAGATGAAACATACAGATTTTATCATCATCAG 1440
Db 1456 TCATCCCAATTCACAGCTGATCTTATTTGCAATGACATACCGATTTTATCATCATCAAGT 1515
Qy 1441 ACTTTCAGAAATTTGACAGAAAGCAAGCACTGTTGTCATATACAGAGCCACATGCT 1500
Db 1516 ACTTTCAGAGATTTGCTGCGCAAGAAATPACATTTGGAACAAATGAGATCACAAGCT 1575
Qy 1501 TTGATCTTCTGCTCTTCAACCGTGTGATCATGATATCATGATGTTGATATCCCAATTC 1560
Db 1576 TTGATCTTCTGCGGCTCTTCAACCGATGTGATCAGAAATGATGTTGATATCCCAAGTTC 1635
Qy 1561 AACATGTTTCCCTGATGATGATATATATCTTCCCTTACACCGAAGAGAGCGG 1620
Db 1636 AACATGTTTCACTGAGAGACATGATATCTTCCCTTACACAGAGCCGATGAG 1695
Qy 1621 AGGTTGAAGCATTTTCAATCTGATGATGAGACCTTCTTACACCAAGTTGAGATGAA 1680
Db 1696 AGGCTGATCTTCTTCAACCTGAAATGAGAGGCTTCTTATATGAAAGTTGAGAAATGAG 1755
Qy 1681 GAACATTAATGCTGCTCATGACCGCAACAGCAATCTGTTTCAACATGCCAAGCTT 1740
Db 1756 GAACATTTGTTGATTTGAAAGATCGGTCCAGGCAATATTCACAAATGCGCAGATTTG 1815
Qy 1741 GATGCTGCAAGAACTTAAACCGGACTGTCGAGTGTGCGGCAAGAACCCAAAGTTGCGT 1800
Db 1816 GATGCTGCAAGAACTTAAACCGGACTGTCGAGTGTGCGGCAAGAACCCGCTTAAAG 1875
Qy 1801 GAGTTGCTTAACTCTGATGTTGATGAGTGTGATAGCGGAAAGAAATCTTAAAGATTTGAA 1860
Db 1876 GAGTTGCTTAACTCTGATGTTGATGAGTGTGATAGCGGAAAGAAATCTTAAAGATTTGAA 1935
Qy 1861 GAGAAAGCTGAAATGAAAGAAATTTTGAAGTATCCAGATTCACAACTTGAAGGCCAA 1920
Db 1936 GAGAAAGCTGAAATGAAAGAAATTTTGAAGTATCCAGATTCACAACTTGAAGGCCAA 1995
Qy 1921 TTCAGATGATATCATCTCAAAATGAACAGAAATCCGAATGTTGAATCTTACGATACAT 1980
Db 1996 TTTGATGATTTTTCGACAGATATACCGATTAAGAAACGATGAATCTACCGATGATTT 2055
Qy 2056 TGCGACAGAAAGGCTTGTGTAACAGCTGATGATGATGAAAGCTTGTGAATGACAGTT 2040
Db 2041 GTGAGGCAATGATTTGCGGTTTTCGCAACATTTGCGCAACTGTAAACGTTGACAGCCGAG 2100
Qy 2116 GTGAGGCAATGATTTGCGGTTTTCGCAACATTTGCGCAACTGTAAACGTTGACAGCCGAG 2175
Qy 2101 ATTATGTCATGGAATCTGTTTCAACATTTGATCTTACATGATGATCAAGTCTGCT 2160
Db 2176 ATTATGTCATGGAATCTGTTTCAACATTTGATCTTACATGATGATCAAGTCTGCTGCT 2225
Qy 2161 GACATCTGTCGATTTCTTGAAGAAAGTATGAAAGATCCATCTCACTGGATTAAGATC 2220
Db 2236 GAATATCTTTCGATCTTCTGAGAAAGGCAAGGAAACCATCATCAATTTGGGATTAAGATC 2295
Qy 2221 TCCCAAGAGGCTTGAAGCAATGAGAGAAAGTATACATGGAAGATTTTACTCGGAGAGA 2280
Db 2296 TCCCAAGAGGCTTGAAGCAATGAGAGAAAGTATACATGGAAGATTTTACTCGGAGAGA 2355
Qy 2281 CTATTTGACCTGACAGAGATGATGATTTGGAAGATGTTTCAACCTTGAAGCGCGT 2340
Db 2356 CTGCTTAACTTAACTGAGATTTTACGATTTTGAAGATGATATCAACCTTGAAGCGCGT 2415
Qy 2341 GAGAGTGTGCTTAACTTGAAGATGTTTATGCTTAAAGTAAACGTTAACTGAGTAAATCA 2400
Db 2416 GAGAGCGCGGTTATTTGAAATGTTTATGAGCTTCAAGTAAACGCAAGCTGAGATTTCT 2475

[illegible]

Db	127	ACTGCTAAATGGAATGAAAATTTAGCTCTTTCAAGAAATTGAACGAAGGAAAAGGA	186
Qy	121	ATTCTGCAACACCATCAAAATATTCTAGAGTTGGAAGCTATCCCTGAAGAGAACGAAG	180
Db	187	ATTTTGGACAACCATCAAGATGATTCGATTTTGGAAAATTTCTTGAAGGAATGACAG	246
Qy	181	AAGCTGCTAATGAGTGCATTTTTTGAAGTATTTGAAGCTATGTCAGAAAGGATGTGTG	240
Db	247	AAGCTGACTAATGAGTGCATTTGGTGAATTTCTCAGATTCACACAGGAAGCTATAGTTTG	306
Qy	241	CCTCCATGGGTTGCATCTGCTGTTCGTCAAAGGCTGTGTTTGGAGTACATTAGAGTG	300
Db	307	CCACCAATGGGTGTGCTGTGCTGTTCGTCAAAGGCAGGTGTCTGGAGTATCTGAGAGTG	366
Qy	301	AATGTTACGCGCCTTGTGTGTTGAGAACTCACTGTGCTGAGTATCTCCACTTCAAGAA	360
Db	367	AATGTGCATCTCTTGTGTGTTGAAAATTTTCCAACTGCTGATGTTTCTCAAAATTCAGAA	426
Qy	361	GAGTTGTATGGAAGTCAAAATGGAATCTTTGTTTGGAAATTTGATTTTGGCCCTTC	420
Db	427	GAACTGTATGGAAGTCAAAATGGAATCTTTGTTGCTTGAATTTGACTTTGAACCAATTT	486
Qy	421	AACTCATCATTTCCCGCCCGCACTCTTCAAAAATCCATTGTAATGATGTGAGTTCTTA	480
Db	487	ACGCAACTTTTCCCTCCTCTTACTCTCAACAATTCATTGGAATGATGTTCAGTTTCTC	546
Qy	481	AATGTCACCTTTTGGGAAATTTGTTCCATGACAGAGAGCATGCCCTTTGCTCGAA	540
Db	547	AACCGTACCTTTTCCGCAAACTCTTCCATGACAGAGAGATTTGTCATCCACTTTTGGAA	606
Qy	541	TTTCTGAGATCCATTTGTCAAGAGGCAABAACATGATGTTGAAATGACAAATTCABAAC	600
Db	607	TTTTCGACTTCACAGCTACAAAGGGGAAACAATTGATTTGAATGACAGAAATTCABAAC	666
Qy	601	TTGAAATGCTCTTCAACATGTTTGGAGAAAGCAGAGAGATCTTGTATCCCTACTCTCT	660
Db	667	CTGATTCCTTTCAACAATGTTCTGAGAAAGCTGGAAGATATCTAGGCACAGTTGCTCCG	726
Qy	661	GAGACACCATGTGCGCAATTCGAACACCGGTTCCAGAAATCGGTTTGGAAAGAGTTGG	720
Db	727	GATACACCGTAATCTCGAATTCGAACACAGGTTCCAGAGATTTGGTTGGAGAGAGTTGG	786
Qy	721	GGTGCACCGCCAGAACCGCGTGTCCGAGATATCCAACTCCTTTTGGATCTTCTTGAAGCA	780
Db	787	GGAGACATGCAAGAGCGCGTGTCCGATTCATTCAGCTTATTTGATCTTCTTGAAGCT	846
Qy	781	ACGATATCTTGCACCCCTTGAAGAGTTCTTGGGAGAAATCCCAATGAGTTCAATGTTGTG	840
Db	847	CTGATATCTTGTACCTCTTGAGACTTTCCTTGAACAGAAATCCCTATGATGTTTAATGTTGT	906
Qy	841	ATTCTCACTCCCAAGGATATCTTCGCTCAAGACAATGTTTGGGGTATCCCGACACCGGT	900
Db	907	ATTCTTTCTCTCATGAGTTACTTTGTCTCAAGATGATGTCTTGGGATATCCCGATACCGGT	966
Qy	901	GGGAGAGTTGTTTATCTTGGATCAAGTCCGAGCTTTGGAGATGAGATGCTCCTCCGT	960
Db	967	GGTCAAGGTTTATCAATTTTGAATCAAGTCTGTGCTTGAAGATGAGATGCTCAATGCG	1026
Qy	961	ATTAAGACAAGAGACTCAACATCAACCCCTCGAATCCTCATTTATCTAGACTTCTCTCT	1020
Db	1027	ATTATAGAAAACAAGGCTTGAATATGCTTCTTCGCAATTTCACTATCACTCGCTTTCTCCA	1086
Qy	1021	GATGCTGTCCGAAACAACATGCGGTCAACGACTTGAGAAAGTATACGGAACAGAGCACTCG	1080
Db	1087	GAGCAGTCCGAACTATGTTGGCAACGACTCGAAGAAAGTCTATGAAACGAGCATTTGC	1146
Qy	1081	GATATTTCTGAGTACCTTCAAGACAGAAAGGAAATTTGTTGAAAATGAGATCTCAAGA	1140
Db	1147	CACATTTCTGAGTGTTCCTTCAAGATCAGAAAGGAAATTTTTCGCAAGTGAATCTCCGCT	1206
Qy	1141	TTTTGAAAANAATCTGGCATATCTTGGAAACCTTACAGAGGATGTTGCTCATGAAATCTCC	1200
Db	1207	TTGG---AATGTGGCCATATCTAGAAACCTTACACGAGATGTTGCTCATGAGCTTGC	1263

QY 1201 AAAGATTGACGGCAGCCAGATCTGATCATCGAAACNACAGCGCATATGCTC 1260
 DB 1264 ABAAGTTGACAGGACCAACGATCTGATGTTGGAAACATGATGAGAAACATGTT 1323
 QY 1261 GCCTCTGCTCGACATAAATAGGTGTCACAGTGCACCATGCCCATGCTTTGAG 1320
 DB 1324 GCTTCTTGTGGACATAAATAGGTGTCACAGTGTGCTATGCTCATGACCTTGAG 1383
 QY 1321 AAGCAAAATATCCAGATTCGATATCTATTTGAAAGAGCTTGAAGCAATATTC 1380
 DB 1384 AAGCATAGTATCTGATCTGATATTTTACGAAAAAATTCGAAGAAATATCACCTC 1443
 QY 1381 TCTTGCCCAATTTACAGTGTATCTTTTTCGATGAAACCATACAGATTCATCATCAG 1440
 DB 1444 TCTTGCCCAATTTACAGTGTATCTTTTTCGATGAAACCATACAGATTCATCATCAG 1503
 QY 1441 ACTTTCAGGAAATTCGAGAGAGAGACATGTTGTCATACGAGAGCCACATGCT 1500
 DB 1504 ACCTTTCAAGAGATTTGCTGGAAGAGATATCTGTTGACATGATGAGATCACATGCC 1563
 QY 1501 TTCACTCTTCTGCTCTACCGTGTGTCATGATGATGATGATGATGATGATGATG 1560
 DB 1564 TTCACTCTTCTGCTCTACCGTGTGTCATGATGATGATGATGATGATGATGATG 1623
 QY 1561 AACATTTGTTCCCTGCTGCTGATATGAGATATATCTTCCCTTACCGGAGAGAGGCG 1620
 DB 1624 AACATTTGTTCCCTGCTGCTGATATGAGATATATCTTCCCTTACCGGAGAGAGGCG 1683
 QY 1621 AGTTTGAAGCATTTTCATCTGATATGAGATGAGATGAGATGAGATGAGATGAG 1680
 DB 1684 CGGTTCATCATTTCTACCGTGAATGAGATGAGATGAGATGAGATGAGATGAGAT 1743
 QY 1681 GAACATTTATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1740
 DB 1744 GAACATTTATGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1803
 QY 1741 GATGCTGTCAGGATTTAACCGGATGCTGCTGATGATGATGATGATGATGATGATG 1800
 DB 1804 GATGCTGTCAGGATTTAACCGGATGCTGCTGATGATGATGATGATGATGATGATG 1863
 QY 1801 GATGCTGTCAGGATTTAACCGGATGCTGCTGATGATGATGATGATGATGATGATG 1860
 DB 1864 GATGCTGTCAGGATTTAACCGGATGCTGCTGATGATGATGATGATGATGATGATG 1923
 QY 1861 GAGAAAGCTGAATGAGAAATGTTGAGCTGATGCAAGTACCACTTGAACGCGCAA 1920
 DB 1924 GAGAAAGCTGAATGAGAAATGTTGAGCTGATGCAAGTACCACTTGAATGCGCAA 1983
 QY 1921 TTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980
 DB 1984 TTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2043
 QY 1981 TGCGACACGAAAGAGCTGCTTGTACAGCTGATGATGATGATGATGATGATGATG 2040
 DB 2044 TGCGACACGAAAGAGCTGCTTGTACAGCTGATGATGATGATGATGATGATGATG 2103
 QY 2041 GTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
 DB 2104 GTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2163
 QY 2101 ATTATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
 DB 2164 ATTATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2223
 QY 2161 GACATATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
 DB 2224 GATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2283
 QY 2221 TCCCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
 DB 2284 TCCCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2343

QY 2281 CTATTAACCTGACAGAGTGTATGATGATGATGATGATGATGATGATGATGATG 2340
 DB 2344 CTATTAACCTGACAGAGTGTATGATGATGATGATGATGATGATGATGATGATG 2403
 QY 2341 GAGAGTGTGTTACCTTGAAGATGATGATGATGATGATGATGATGATGATGATG 2400
 DB 2404 GAGAGTGTGTTACCTTGAAGATGATGATGATGATGATGATGATGATGATGATG 2463
 QY 2401 GTTCATTGGACAGAGTGA 2420
 DB 2464 GTTCATTGGACAGAGTGA 2483

RESULT 11
 VSUSCS
 LOCUS 2647 bp mRNA linear PLN 11-MAY-1995
 DEFINITION V. faba mRNA for sucrose synthase.
 ACCESSION X69773
 VERSION X69773.1 GI:22037
 KEYWORDS starch synthetase; sucrose synthase.
 SOURCE
 ORGANISM
 Vicia faba
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae;
 Vicia.
 1 (bases 1 to 2647)
 REFERENCE
 AUTHORS Heim, U., Weber, H., Baumlein, H. and Mobs, U.
 TITLE A sucrose-synthase gene of Vicia faba L.: expression patterns in
 developing seeds in relation to starch synthesis and metabolic
 regulation
 JOURNAL Planta 3, 394-401 (1993)
 REFERENCE
 AUTHORS Heim, U.
 TITLE Direct Submission
 JOURNAL Submitted (11-DEC-1992) U. Heim, Institute of Plant Genetics & Crop
 Plant Research, Corrensstr. 3, 0-4325 Gatersleben, FRG

FEATURES
 source
 1. .2647
 /organism="Vicia faba"
 /mol_type="mRNA"
 /cultivar="Frio"
 /db_xref="taxon:3906"
 /clone="pubx11/20/2"
 /clone_lib="CDNA library"
 /pop_variant="minor"
 1. .2647
 /gene="VfSuc"
 <1. .2647
 /gene="VfSuc"
 34. .2454
 /gene="VfSuc"
 /BC number="2.4.1.13"
 /codon_start=1
 /product="sucrose synthase"
 /protein_id="CAA9428.1"
 /db_xref="GI:22038"
 /db_xref="GI:22038"
 /db_xref="SWISS-PROT:P31926"
 /translation="MATERLTAVSLERDELTLTANRNLALLSRIEAKGKILGH
 HOYIAEBEIPENROKLTGAFGEVIVRSGEIVLPEPVALVPRPGVWEYLVAV
 HALVVENTLOPAEFLPKKEELVDGSANRNPFLIEDPFTASPRPLTKRAGVQL
 NRHLISAKLPHDKESLHLEFLPKLHSHKGTLLMNDRIQPDSDIQLHRAEELSLV
 DPEPTVSEFERHREIGLERGKDSARVLESIQLLDLLEADPECTLETFDLRIEIV
 FNVVILSPHGYFADVDVLYGPDGGOVYLLDVRALSEMLNRIRKQGLDIVRLLI
 ITRLLPDAVGTGGRLEKYGTGHECHILLVPRDQGIYRKVLSREVPYLETYTE
 DVABELAKELQKPDILVNGSDNIVASLLAHKLGVTCTIAHALEKYPESDIYE
 KKEBEKYESCOPTADLFAHHNDPILITSTFOIAGSKDITVGOVESHATFLIGLVYV
 VHGIIDVDPKENVISRGADDTIYFPYTESRRLTSFPPELILYXSVNBEHICVK
 DRSEPLITFVARLDKRVNITGLVEMYGKAKLBEVLVAVVAGDRKESDGLTEKAM
 KMYBELLETYKLGQFWMISQNRVNGELVYVCTKGFVQPAVAFGLTEKAM
 MANGLPFATLNGPAPAIIVHGSGFPIIDYHGRADLLVERFEKVKADPSHDKIS
 LGGLRIEBEKYTIQYSQRLTLTLGVYGFVKHVSINDRLBSRRYLEMFAVKRKLAE

QY 2101 ATTATGTCATGGGAATCTGTTCAACATTGATCTTACCAATGATGATCAAGTCT 2160
DB 2131 ATCATGTCATGGAATCTGATGATCAATGATGATCAATGATGATGATGATGAT 2130
QY 2161 GACATGTCATGATCTTCTTGAAGAGTGAAGAAAGTCAATCTGATGATGATGAT 2220
DB 2191 GATCTCTGTCATGATCTTCTGAGAAAGTGAAGGATGATGATGATGATGATGAT 2250
QY 2221 TCCCAAGAGGCTGGAAGCAATGAGAGAGATGATGATGATGATGATGATGATGAT 2280
DB 2251 TCTCTGAGGCTGCAACGATGATGAGAGATGATGATGATGATGATGATGATGAT 2310
QY 2281 CTATTGACCTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
DB 2311 CTCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2370
QY 2341 GAGAGTGTGCTGATCTTGAAGATGATGATGATGATGATGATGATGATGATGAT 2400
DB 2371 GAGAGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2430
QY 2401 GTTCCATGTCAGAGAGAT 2420
DB 2431 GTGCTCTGAGCTGTTGAGA 2450

RESULT 12
VFAUDPFTTA 2665 bp mRNA linear PLN 23-SEP-1993
LOCUS Vicia faba UDP-glucose:D-fructose-2-glucosyltransferase mRNA,
DEFINITION complete cds.
ACCESSION M97551.1 GI:295425
VERSION M97551
KEYWORDS UDP-glucose:D-fructose-2-glucosyltransferase.
SOURCE Vicia faba (fava bean)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae;
Vicia.
REFERENCE 1 (bases 1 to 2665)
AUTHORS Kuster,H., Fruhling,M., Perlick,A.M. and Puhler,A.
TITLE The sucrose synthase gene is predominantly expressed in the root
nodule tissue of Vicia faba
JOURNAL Mol. Plant Microbe Interact. 6 (4), 507-514 (1993)
MEDLINE 84003420
PUBMED 8400379
COMMENT Original source text: Vicia faba (strain Kleine Thueringer) root
nodule cDNA to mRNA.
FEATURES
source location/Qualifiers
1..2665
/organism="Vicia faba"
/mol_type="mRNA"
/strain="Kleine Thueringer"
/db_xref="taxon:3906"
/cfeature_type="root module"
1..40
/note="putative"
41..2461
/EC_number="2.4.1.13"
/function="cleavage of sucrose"
/standard_name="sucrose synthase"
/note="putative"
/codon_start=1
/product="UDP-glucose:D-fructose-2-glucosyltransferase"
/protein_id="AAC37346.1"
/db_xref="GI:295425"
translation="MATERLTVHSLSIRELDELTLNANREILALISRIEAKGILGH
HQAVALREPEREENROKLTGAPGVALVSPQEAIVLPWVALAVRPGWVEYLRNV
NALVNTQAPAEFLKFEELVDGANGFVLEDPPTASFPPTLANKSIGKVOEL
HRHLSAKLFDKESLHPLFLRLHSYKGLMLNDRIQNDLSQHLVRAKEELSTV
DPEYVSEFERFOEIGLBERGMSABVESIQLLDLLEAPDCTLEPLDRIIPWV
FNVILSPHYFAODVLYGYDGTGGVYVLDQVRALESBMLNRIKKGGDIDVIRILI

ITRLPDVAVITGQRLKXYGTEHCHILRVPEDOKIVKMWISREWPYLETYE
DVHLEAKELQKRPDLIVGNVSDGNIIVASLIHLAQTOCIATLAKTKPSESDIY
KKEPKKHFSCQFADLIPANMHDPIITSDYHAGSKDVTGQYSEHAPLPLXYV
VHGIDVDPKPMIVISPGADQTIYPRYETSRRLTSFVPIBELIYSTVNEBHI CVLK
DRSEPIITPMARLDVRNKTIGVBYWGRKAKRELVLVYVADGRKRSKDLERKAE
KRWELIETKILNGQFRWISSQMRVNGELVRICTKFAFVQPAVEAFGLVEA
MARGLPFATLNGPARIIVHKGSGFIIDYHGRADLDLVEFEKKAQPSHWDKS
LGGIQRLEBKTYQMIVSQRLITLVGVGFWMKHSNLDRLRESRYLMEFYALKYKRLAE
SVPLAVEE"
2462..2665
/note="putative"
2529..2534
3'UTR
orig polyA_signal
Query Match 58.6%; Score 1539; DB 8; Length 2665;
Best Local Similarity 77.5%; Pred. No. 0;
Matches 1876; Conservative 0; Mismatches 541; Indels 3; Gaps 1;
QY 1 ATGAGTGAAGCGTGTCTCACTGCGCTCCACAGTCTCCGTGAGCGTTGGATGAGACCTT 60
DB 41 ATGAGTGAAGCGTGTCTCACTGCGCTCCACAGTCTCCGTGAGCGTTGGATGAGACCTT 100
QY 61 CTGCTCAAGAAAGAGATTTGGCTGCTCTCAAGATCGAGGCAAGAAAGAA 120
DB 101 ACTGTAATGAGATGAAATTTAGCTCTTATCAAGATGAAAGCAAGGAAAGGGG 160
QY 121 ATTCTGAACCACTAAATTAATCTAGATGATGAAAGTAAAGTAAAGTAAAGTAAAG 180
DB 161 ATTTTGACATCAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 220
QY 181 AAGCTGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 221 AAGCTGCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 280
QY 241 CTTCATGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 281 CCACCAATGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 340
QY 301 AATGTTACCGCTTGTGTTGAGAACTGATGATGATGATGATGATGATGATGATGATGAT 360
DB 341 AATGTTACCGCTTGTGTTGAGAACTGATGATGATGATGATGATGATGATGATGATGAT 400
QY 361 GAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 401 GAATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 460
QY 421 AACTCATATTCCTCCGCCCACTCTTCAAAATTCATGATGATGATGATGATGATGATGAT 480
DB 461 ACGGCTGCTTCTCCCTGCTCACTCTCAACAGTCAATGATGATGATGATGATGATGATGAT 520
QY 481 AATGTTACCGCTTGTGTTGAGAACTGATGATGATGATGATGATGATGATGATGATGAT 540
DB 521 AATGTTACCGCTTGTGTTGAGAACTGATGATGATGATGATGATGATGATGATGATGAT 580
QY 541 TTCTCTGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 581 TTCTCTGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 640
QY 601 TTGAATGCTTCAACATGTTTGAAGAAAGAGAGATGATGATGATGATGATGATGATGAT 660
DB 641 CCTGATTCCTTCAACATGTTTGAAGAAAGAGATGATGATGATGATGATGATGATGATGAT 700
QY 661 GAGACACAGTGGCGAATTCGACACCGGTCCAGAGAAATCGGTTGGAAAGAGGTGG 720
DB 701 GAAATCTCGTACTCGAATTTGAACACAGGTTCCAGAGAAATCGGTTGGAAAGAGGTGG 760
QY 721 GGTGACACCGGAGACCGGTGCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 761 GGAACACCGGAGACCGGTGCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 820
QY 781 ACTGATCTTGAACCTTGAAGAGTCTCTTGAAGAAATCCCATGATGATGATGATGATGAT 840
DB 821 CTGATCTTGAACCTTGAAGAGTCTCTTGAAGAAATCCCATGATGATGATGATGATGATGAT 880

QY 841 ATTCTCACTCCCAACGAGATCTTCCTCAAGACAATGTTTGGGTATCCGACACCGGT 900
 DB 881 ATTCTTCTCTCTCAATGAGTTACTTGTCTCAAGATGATGCTTGGGATACCTGATACGTGT 940
 QY 901 GGGCAGGTGTTTACATCTTGGATCAAGTCCGAGCTTGGAGATGAGATGCTCTCCGT 960
 DB 941 GGTCAAGGTGTTTACATCTTGGATCAAGTCCGAGCTTGGAGATGAGATGCTCTCCGT 1000
 QY 961 ATTAAGCAACAGAGACTCAATCAACCCCTGGAATCCCTCATTTACTAGACTTCTCTCT 1020
 DB 1001 ATTAAAGAAACAGGCTTGGATATGCTTCTGCAATCTCATTTACATCTGCTCTCTCTCT 1060
 QY 1021 GATCTGTGAGAACATATGCGGTCAACGATTTGAGAAAGTATACGGAACAGACACTCG 1080
 DB 1061 GATCAGTCCGAACTACTGTGGCCAAAGCTTGAAGAAAGTCTATGGAACCGACATGTGT 1120
 QY 1081 GATATTTCTGAGATACCTTTCAAGAACAGAAAGGAAATTTGTCGAAATGATCTCAAGA 1140
 DB 1121 CACATTTCTGAGTCCCTTCAGAGATCAGAAAGGAAATTTGTCGAAATGATCTCAAGT 1180
 QY 1141 TTTGAAAAGCTGGCCATCTTGGAAACCTACAGAGATGTTGCTCATGAATCTCC 1200
 DB 1181 TTGG--AAGTCTGACATATCTAGAAACCTACAGAGATGTTGCTCATGAGCTTGGC 1237
 QY 1201 AAGAGTTGACGCGACGCGAGATCTGATCATCGGAAACNACAGCGACGCAATATGCTC 1260
 DB 1238 AAGAGTTGACGCGAAACGAGATCTGATTTGTTGAAACCTACAGATGAGAAACATTTGT 1297
 QY 1261 GCTCTCTGCTGCAATTAATGAGTGTCAACAGTGCACATCGCCATGCTTTGGAG 1320
 DB 1298 GCTCTTGTGTGCAATTAATGAGTGTCTCACTGATCTATGCTCATGACCTTGGAG 1357
 QY 1321 AAGCAAAATATCCAGATTCAGATATCTATTTGAAGAACTTGAAGCAATATCATTTTC 1380
 DB 1358 AAGACCAAGTATCTGAATCTGATATTTACGGAAGAAATTTGAAGCAATATCATTTTC 1417
 QY 1381 TCTTGCCAAATTAAGCTGATCTTTTTCATGAAACATATGATTCATCATCAAGT 1440
 DB 1418 TCGTGCCAAATTAAGCTGATCTTTTTCGCAATGAAACATGATTCATCATCAAGT 1477
 QY 1441 ACTTTCAGAAATTCAGAGAACAGACACTGTTGCTCAATACAGAGCCCACTGCT 1500
 DB 1478 ACTTTCAGAGATTCGAGAACAGACACTGTTGCTCAATACAGAGCCCACTGCT 1537
 QY 1501 TTCACTCTCTGCTGCTACCGTGTGTACATGATGATGATGATGATGATGATGATGATGAT 1560
 DB 1538 TTCACTCTCTGCTGCTACCGTGTGTGTGACGATGATGATGATGATGATGATGATGATGAT 1597
 QY 1561 AACATTTGTTCCCTGCTGCTGATATGAGATATATCTTCCCTTACACCGAAGAGAGCGG 1620
 DB 1598 AACATTTGCTTCCAGAGAGCTGATACAGACATTTACTTCCCTTACACCGAAGAGAGCGG 1657
 QY 1621 AGTTGAAGCAATTCATCTGAGATGAGACACTTCTTTACACCAAGATTTGAGATGAA 1680
 DB 1658 AGTTGAAGCAATTCATCTGAGATGAGACACTTCTTTACAGACAGTGAAGAAATGAA 1717
 QY 1681 GAACCTATATGCTGCTCAATGACCGCAACAGCAATTTCTGTCAGAAATGCCAAGGCTT 1740
 DB 1718 GAGCAGATATGCTGCTCAATGACCGCAACAGCAATTTCTTACCATGCGAGGCTT 1777
 QY 1741 GATCGTGTCAAGAACTTAACCGGACCTGCTGAGTGTGCGGCAAGAACCCAAAGTTGCT 1800
 DB 1778 GATCGTGTCAAGAACTTAACCGGACCTGCTGAGTGTGCGGCAAGAACCCAAAGTTGCT 1837
 QY 1801 GAGTTGCTAACTCTGATGTTAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGT 1860
 DB 1838 GAGTTGCTAACTCTGATGTTAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGT 1897
 QY 1861 GAGAAAGCTGAATGAAGAAATGTTTGAAGTGTGATGAGTGTGATGAGTGTGATGAGTGT 1920
 DB 1898 GAGAAAGCTGAATGAAGAAATGTTTGAAGTGTGATGAGTGTGATGAGTGTGATGAGTGT 1957

QY 1921 TTCAATGATATCATCTCAATGAACAGAAATCCGAAATGTTAACTTTACCGATACATT 1980
 DB 1958 TTCAATGATATTTCTCTCAGATGAACCGGTGTCAAGAAATGGAAGCTTACCGGTATATC 2017
 QY 1981 TGGACACGAAAGAGTCTTGTGACAGCCCTGATGATGATGATGATGATGATGATGATGATGAT 2040
 DB 2018 TGGACACGAAAGAGTCTTGTGACAGCCCTGATGATGATGATGATGATGATGATGATGATGAT 2077
 QY 2041 GTGAGCAATGACTTGCCTGTTTGCACAACTTTCGAAACCTGTAAACGCTGAGACGCGAG 2100
 DB 2078 GTTGAAGCCATGAGCCATGAGATTTACCAATTTGCAACATGATGATGATGATGATGATGAT 2137
 QY 2101 ATTAATGCTGATGAAATTTGCTTCAACATGATGATGATGATGATGATGATGATGATGATGAT 2160
 DB 2138 ATCAATGCTGATGAAATTTGCTTCAACATGATGATGATGATGATGATGATGATGATGATGAT 2197
 QY 2161 GACATCTGCTGATTTCTTTGAAAGTGTGAAAGATCCATCTCACTGGATTAAGATC 2220
 DB 2198 GATCTCTGCTGATTTCTTTGAAAGTGTGAAAGATCCATCTCACTGGATTAAGATC 2257
 QY 2221 TCCCAAGAGCTTGAACGAAATGAGAGAGATATACATGAAAGATTTTACTCGAGAGA 2280
 DB 2258 TCTCTGCTGCTGCTTCCAGATATTTGAAGAGATGAGAGATTTTACTCGAGAGA 2317
 QY 2281 CTATGACCTTGAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
 DB 2318 CTCTTACCTCACTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2377
 QY 2341 GAGAGCTGCTGATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
 DB 2378 GAGAGCTGCTGATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2437
 QY 2401 GTTCAATGCGAGAGAGATA 2420
 DB 2438 GTGCTCTACCTGTTGAGGA 2457

RESULT 13

BD236049 2906 bp DNA linear PAT 17-JUL-2003

LOCUS

BD236049 Materials and method for modification of plant cell wall

DEFINITION

polysaccharides.

ACCESSION

BD236049 BD236049.1 GI:33045819

VERSION

JP 2002527056-A/71.

KEYWORDS

Bucalyptus grandis

SOURCE

Bucalyptus grandis

ORGANISM

Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE

1. (bases 1 to 2906) Bloksberg, L.N.

AUTHORS

Materials and method for modification of plant cell wall

TITLE

PATENT: JP 2002527056-A 71 27-AUG-2002; FLETCHER CHALLENGE

JOURNAL

GENESTIS RESEARCH AND DEVELOPMENT CORP LTD, FLETCHER CHALLENGE

COMMENT

OS Eucalyptus grandis (flooded gum)

PN

JP 2002527056-A/71

PD

27-AUG-2002

PF

08-OCT-1999 JP 2000575985

PR

13-OCT-1998 US 09/170862, 11-AUG-1999 US 60/148426 P1

PC

LEONARD NATHAN BLOKSBERG C12N15/09, A01H5/00, C12N5/10, C12N9/00, C12N9/12, C12N15/00, C12N5/00

CC

Materials and method for modification of plant cell wall CC

FH

Key polysaccharides

FT

Location/Qualifiers

FT

Location/Qualifiers

FEATURES

1..2906 location/Qualifiers

source

location/Qualifiers

FT

Location/Qualifiers

FT

Location/Qualifiers

FEATURES

1..2906 location/Qualifiers

source

location/Qualifiers

FT

Location/Qualifiers

FT

Location/Qualifiers

FEATURES

1..2906 location/Qualifiers

source

location/Qualifiers

FT

Location/Qualifiers

FT

Location/Qualifiers

ORIGIN

/db_xref="taxon:71139"

Query Match 57.1%; Score 1499.4; DB 6; Length 2906;
Best Local Similarity 76.5%; Fred. No. 0;
Matches 1852; Conservative 0; Mismatches 567; Indels 3; Gaps 1;

1 ANGGCTGACCGGCTCTCACTCGAGTCCAGAGTCCCGTGAAGCGTTGGATGAGACCTT 60
223 ATGGCTGATCCGATCTCACTCGAGTCCAGAGTCCCGTGAAGCGTTGGATGAGACCTT 292
QY 61 CTTCGCTCAGAGAGAGATTTGGCCCTTCAGAGTCCAGAGTCCCGTGAAGCGTTGG 120
DB 293 TCTGCTCAGAGAGAGATTTGGCCCTTCAGAGTCCAGAGTCCCGTGAAGCGTTGG 352
QY 121 ATTCTGCAACACCATCAATTTATTCTAGAGTTGAAGTATCCCTGAGAGAGAGAG 180
DB 353 ATCTTGACGCGCCACACAGATTTTGGCTGAGAGTTCAGAGTCCCGTGAAGCGT 412
QY 181 AAGCTGCTAAGAGAGATTTTGGAGTGAAGTGAAGGCTGAGTCAAGAGCGATGCTG 240
DB 413 AAGCTTCTGATGAGAGGCTTGGTGAAGTCTCAATCACTCAGAGAGCGATGCTG 472
QY 241 CCTCAGAGGAGTGAAGTCTGCTGCTGCTCAAGGCGCTGCTGCTGAGTCAATGAGT 300
DB 473 CCTCAGAGGAGTGAAGTCTGCTGCTGCTCAAGGCGCGGCGTGGAGACATCGCTG 532
QY 301 AATGTTCAAGCGCTTGTGTTGAGAGTCACTGCTGCTGAGTATCTCACTTCAAGAA 360
DB 533 AAGCTGCAAGCGCTTGTGTTGAGAGTCACTGCTGCTGAGTATCTCACTTCAAGAA 592
QY 361 GAGCTGTTGAGAGAGTCAATGAGAACTTGTGTTGAGTGAAGTGAAGCGCTT 420
DB 593 GAGCTGTTGAGAGAGTCAATGAGAACTTGTGTTGAGTGAAGTGAAGCGCTT 652
QY 421 AACTCATATTTCCCGCCGCAACTCTTCAAAATCATTTGATGATGAGAGTTCCTA 480
DB 653 AACTGCTTTTTCGCGCCGCAACTCTTCAAAATCATTTGATGATGAGAGTTCCT 712
QY 481 AATGTCACCTTTGCGGCAAAATTTGTTCAATGAGAGAGAGAGTCACTTGTCTGAA 540
DB 713 AATGTCACCTTTGCGGCAAAATTTGTTCAATGAGAGAGAGTCACTTGTCTGAA 772
QY 541 TTCTCAGAGTCAATGTCAGAGGCAAGAGATGATGAGTGAAGTGAAGTCAAGAC 600
DB 773 TTCTCAGAGTCAATGTCAGAGGCAAGAGATGATGAGTGAAGTGAAGTCAAGAC 832
QY 601 TTGAATGCTCTTCAACATGTTTGAAGAGAGAGAGTATCTTGTATACCTTACT 660
DB 833 GTGTTCTCTTCAACATGTTTGAAGAGAGAGAGTATCTTGTATACCTTACT 892
QY 661 GAGACACATGTCGCGAATTTGAACAACCGGTTCCAGAAATCGGTTGAAAAGGTTGG 720
DB 893 GAGACACATGTCGCGAATTTGAACAACCGGTTCCAGAAATCGGTTGAAAAGGTTGG 952
QY 721 GGTGACACCGGAGAGAGGCTGAGATGAGTCAATCTTGTGATGATCTTGAAGGCA 780
DB 953 GGTGACACCGGAGAGAGGCTGAGATGAGTCAATCTTGTGATGATCTTGAAGGCT 1012
QY 781 ACTGATCTTCAACCTTGAAGAGTCTTGTGAGAGATCCCATGATGATGATGATG 840
DB 1013 CCGACCGGTCACCTTGAAGAGTCTTGTGAGAGATCCCATGATGATGATGATG 1072
QY 841 ATTCTCACTCCCGAGATGATGCTGCTCAAGACAAATGTTGGGATATCCGACACGG 900
DB 1073 ATTCTCACTCCCGAGATGATGCTGCTCAAGACAAATGTTGGGATATCCGACACGG 1132
QY 901 GGGAGAGTGTGTTCACTTGAATCAAGTCCGAGGTTTGAAGATGAGATGCTCTCT 960
DB 1133 GGGAGAGTGTGTTCACTTGAATCAAGTCCGAGGTTTGAAGATGAGATGCTCTCT 1192
QY 961 ATAAAGCAACAGAGTCAATCAACCTCCCTGAGATCTTATATTAATTAATCTTCT 1020

DB 1193 ATTAAGCAACAGAGTCAATCAACCTCCCTGAGATCTTATTAATCACTCGCTTCTTCA 1252
QY 1021 GATGCTGCGGAGACAACTGCGTCAACGATCTTGAAGATGATGAGAGACGACTCG 1080
DB 1253 GACGCGGTTGAGACACTGCGTCAACGATCTTGAAGATGATGAGAGACGACTCG 1312
QY 1081 GATATTTCTCGAGTACCTTCAAGACAGAAAGGAAATTTGTTGAAATGATCTCA 1140
DB 1313 CACTTCTCGGCTCTTCTTCAAGATGAGAGAGGAGTCTCCGAGATGATTTCCCGG 1372
QY 1141 TTTGAAAAGTCTGCGCATCTTGAACCTTCAACAGAGATGTTGCTCATGAATCTTC 1200
DB 1373 TTTG---AGTGTGCGCTTATTTGAAGATGATGAGATGTTGCGAGGAACTTGTCT 1429
QY 1201 AAGAGTTCAGCGGACGCGATCTGATCTGAGTGAAGTGAAGTGAAGTGAAGT 1260
DB 1430 GAGAGTTCAGCGGACGCGATCTGATCTGAGTGAAGTGAAGTGAAGTGAAGT 1489
QY 1261 GCTCTCTTGTGCGACATTAATTAAGTGTCAACAGTGCACATCGCCATGCTTTGGAG 1320
DB 1490 GCTCTCTTGTGAGACATTAATTAAGTGTCAACAGTGTCAACATGAGTGTGAG 1549
QY 1321 AAGACAAATATCCAGATTCAGATATCTTGAAGAGCTTGAAGACAAATACCATTT 1380
DB 1550 AAGACAAATATCCAGATTCAGATATCTTGAAGAGCTTGAAGACAAATACCATTT 1609
QY 1381 TCTTGCCAAATTAAGTCACTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1440
DB 1610 TCTTGCCAAATTAAGTCACTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1669
QY 1441 ACTTTCAGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 1670 ACTTTCAGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1729
QY 1501 TTGACTCTTCTGCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
DB 1730 TTGACTCTTCTGCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1789
QY 1561 AACTTTGTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
DB 1790 AACTTTGTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1849
QY 1621 AGGTGAGAGATTTCAATCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB 1850 CGGTGAGAGATTTCAATCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1909
QY 1681 GAGACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1910 GAGACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1969
QY 1741 GATGCTGCTCAAGATTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
DB 1970 GATGCTGCTCAAGATTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2029
QY 1801 GAGTGTGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 2030 GAGTGTGCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2089
QY 1861 GAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
DB 2090 GAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2149
QY 1921 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
DB 2150 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2209
QY 1981 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
DB 2210 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2269
QY 2041 GTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
DB 2270 GTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2249

Oy	2101	ATTATTTGTCATGGGAATACTCGTGTTCAACTAATGATCCTTACCAATGGTGATCCAACTGCT	2167
Db	2330	ATCATTTGCACTGGCAAAATCGGGCTACACAATTGATCCTTACCAGTAGTGACCAAGCCGGC	2389
Oy	2161	GACATACTCGTCGATTTCTTTGAAAAGTGTAAGAAGATCCATCTCACATGGGATTAAGATC	2220
Db	2390	GAGCTTCTTTGAGACTTTCTTCAACAAGTGCAGAGATTGACAGTCCCACTGGGACGAGATC	2449
Oy	2221	TCCCAGAAGGCTTGAAAAAGATAGAGAGAAATATACATGGAAATTTACTTCGGAGAGA	2280
Db	2450	TCAAAGSGTGCCATGCAAGATTTGAAGAGAGATTCATGGAATAATATTTGAGAGG	2509
Oy	2281	CTATTGACCCCTGCACAGAGAGTGATGGAATTTCTGGAAGCATGTTTTCCAACCTTGAACGCCGT	2340
Db	2510	CTGTGGAACCTGACTGCGCTGTATGGCTTTCTGGAAGCATGTGACTTAACTTGATCGGCGC	2569
Oy	2341	GAGAGTGGTGGTTACCTTGAGATGTTTTATGTCTTTAAATGACCGTAAGCTGGCTGAATCA	2400
Db	2570	GAGAGTCCCGCGTACTTGGAAATGTTCTATGGCCCTCAAGATATCGCCCACTGGCACAGTCT	2629
Oy	2401	GTTCCATTGGCAGAGGAGTAAA	2422
Db	2630	GTTCTCCGGCTGTGAGTAAA	2651

QY	6	CTTGTCTACAGGAACGAGATTTTGGCCCTGTCTCAAGATTCGAGCGGCAAAAGSAAAAGCA	120
Db	302	TTCTCTACCGCAACGATATGTGGCTCTCTTTCAGAGGTTGAGCCAAAGGCCAAAAGC	361
QY	121	ATTCTGCACACATCAATTAATTTCTAGAGTTTGAAGCTATCCCTGAAAGAAACAGAAAG	180
Db	362	ATCTTGGAGCGCCACAGATTTTGTGCGAGTTTGAGGCAATCTGTAGGAAGACGAGACA	421
QY	181	AAGCTCGTAATAGTGCATTTTGTGAAAGTATGAAAGCTAGTCAGAAAGCAATCGTTTG	240
Db	422	AAGCTCTGTGATGGGGCCTTGGAGAGTCTCTAAATCTACTCAGGAAGCCATTGTGTGCG	481
QY	241	CCTCCATGGGTGTCATTTGCTGTTCGATCGAACAGGCGTGGTGTTTGGAGATACATTAGAGTGG	300
Db	482	CCTCACAAGGTTGCTCTTTCGTCTTTCGTCGACAAAGCCGGCGTGTGGAGCAATCCGTGTG	541
QY	301	AATGTCAAGCCCTGTGTGTGTGAGGAATCACTGTGTGTAGATATCTCACTTCAAGGA	360
Db	542	AACGTTCATGCGCTGTGTCTTGACCAATTTGAGAGTGTGCTGAGTATCTGCACTTCAAGAA	601
QY	361	GAGCTGTGTATGGAAGTTCAATAGSAACTTGTTTTGGAAATTGGATTTGAGCCCTTC	420
Db	602	GAGCTGTGTATGGAAGCTGAATGGATCACTTGTGTGCTGTGAGCTTGACCTTGGAGCATTC	661

RESULT 14	BD235993	2913 bp	linear	PAT 17-JUL-2003
LOCUS	BD235993			
DEFINITION	Materials and method for modification of plant cell wall polysaccharides.			
ACCESSION	BD235993			
VERSION	BD235993.1	GI:33045763		
KEYWORDS	JP 2002527056-A/15.			
SOURCE	Eucalyptus grandis			
ORGANISM	Eucalyptus grandis			
	Euparieta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Myrtales; Myrtaceae; Eucalyptus.			
REFERENCE	1 (bases 1 to 2913)			
AUTHORS	Bloksberg, L.N.			
TITLE	Materials and method for modification of plant cell wall			
JOURNAL	Patent: JP 2002527056-A 15 27-AUG-2002;			

OY	481	AATGTCACCTTTGGCAMAATTGTCATGACAAGGAGACATGACCCCTTGTCTGAA	540
Db	722	AATGSCATCTCTCCGCTAAGCTCTTCATGACAAGGAAAGCTTGACCCCTGCTGAA	781
OY	541	TTCTTCAGATCCATTGTCACAAGGCAAGAACTGATGTTGAATGACAGAAATTGAGAC	600
Db	782	TTCTCTCAAGTCACCTGCTACAMAGGAGAACTGATGGAATGCCAAGATCCAGAT	841
OY	601	TTGATGCTCTTCAACATGTTTGAAGAAAGCAAGAGATATCTTGTAACCTTACCTCT	660
Db	842	GTGTTTCCCTCCCAACATGTCCTGAGGAAGCGGAGATATCTGACCTGCTCAAAACC	901
OY	661	GAGACACCATGTGCGGAATTGAAACCCGGTTCAGAGAAATCGGTTTGAAAGAGGTGG	720
Db	902	GAGACCCCTGATCTCCAGTTCTGACACAAGTTCCAGAGAAATCGGAGCTCGACCGGGGGTGG	961

COMMENT	OS	PN	PD	PF	PR	PC	CC
	<i>Eucaalyptus grandis</i> (flooded gum)	JP 2002527056-A/15	27-AUG-2002	08-OCT-1999 JP 2000575985	09/170862, 11-AUG-1999 US 60/148426 PI	LEONARD NATHAN BLOKSBERG PC C12N15/09, A01H5/00, C12N5/10, C12N9/00, C12N9/12, C12N15/00, C12N5/00	Materials and method for modification of plant cell wall CC polysaccharides
FEATURES							
source							
FT	key						
FT	source						
FT							
	Location/Qualifiers						
	1..2913						
	Location/Qualifiers						
	1..2913						
	/organism="Eucaalyptus grandis"						
	/mol_type="genomic DNA"						
	/db_xref="taxon:71139"						
ORIGIN							

Db 962 GGTACACGGCTGACGGCGCTCTCGAAGTATCCAGCTCTCTTGATCTCTTGAGGCT 1021
 Qy 781 ACTGATCCTTGACACCCCTTGAGAACTCTTGAGAAATCCCATGGTGTCAATGTTGTG 840
 Db 1022 CCCGACCCGTCATCTCTCGAAGAACTTTTGATGAGGTCCATGGTCTTCAACGTCTGT 1081
 Qy 841 ATTCTCACTCCCAACGGATCTTGCTCAAGACAACTGTTTGGGTATCCGACACGGT 900
 Db 1082 ATCATGTCCTCCCAACGGATCTTGCTCAAGACAACTGTTTGGGTATCCGATACGGT 1141
 Qy 901 GGGCAGGTGTTTTCATCTTGATCAAGTCGAGCTTGAGAAATGAGATCTCTCCGT 960
 Db 1142 GGCACGGTGTGTTTCATCTTGATCAAGTGTGTCCTTAGAGAAAGAAATGCTTACCGC 1201
 Qy 961 ATTAAGCAACAAGATCTCAACATCAACCCCTGAAATCTCATATATTACTAGCTTTCT 1020
 Db 1202 ATTAGCAACAAGACTGAGATATTACTCTTCGAGATCTCATTTACCTACCTGGCTTCTTCA 1261

Query March	57.1 %	Score 1497.8	DB 6	Length 2913
Best Local Similarity	76.4 %	Prod. No. 07		
Matches 1851	Conservative 0	Mismatches 568	Indels 3	Gaps 1

Oy	1	ATGGCTGAGCGCTCTCACTCGCGTCACAGCTTCCTCGAGCGCTTTGGATGAGACCCCTT	60
db	242	ATGGCTGATGCGATGTTGATCGACACCAACACCTTGGGAGCGCTTTGGAGAGAACCTTC	301

Accession	Sequence	Position
Dd	1262 GACCCGGTTGGAAACCACTCTGTGGCCAGCGCTTTAGAAAAGTTTGGAGCCGAGTACTCC	1321
Oy	1081 GATATTTCTTCAGTACCTCTCAGAACAGAAAGGAAATTGTTCGAAAAATGATCTCAGA	1140
Dd	1322 CACATTTCTTCGCGCTCCCTTCAGAAATGAGAGGAGATCGTCCGCAAGTGGATTTCGCCG	1381
Oy	1141 TTTGAAAAAGCTCGGCATATCTTGGAAACCTTACACAGAGATGTGTCTCATGAATCTCC	1200

Qy	241	CTCTCAGTGGGTGCACTTGGCTGTCTGCTCCAGAGCCGCGGTGTTTGGGAATCAATTAGACTG	300
Db	687	CTTCATATGGGTGCTCTGTGCTGTCTGCTCCAGAGCCGCGCGGTGGGAACAATCCGTGTG	746
Qy	301	AATGTTCACGCCCTGTGTGTGGAGAACTCACTGTGTGATATCTCACCTTCAAGAA	360
Db	747	AACGTCCATGGCGCTGTGTTCTTGAAGAAATTGGAAGTGTGTGATATCTCACTTCAAGAA	806
Qy	361	GAGCTTGTGTGATGGAAGTTCAATGAAACTTTGTTTGGAAATTGGATTTTGAAGCCCTTC	420
Db	807	GAGCTTGTGTGATGGAAGCTTGAAATGTAACTTTGTGCTTGAAGCTTGACTTTGAGCAATTTC	866
Qy	421	AACATCATATCCCCCGGCCCACTCTTCAAAATCCATTTGGAAATAGTGTGAAGTTTCTTA	480
Db	867	ACTGCTCTTTTTCGCGGCCGCACTTTTCCAAAGTATATTGGCAATGGCGTGAGATTTCCT	926
Qy	481	AATGCTCACCTTTCGGGCAAAATTTGTTCCATGACAAGAGACATGACCCCTTGTGCTCGAA	540
Db	927	AATGCCATCTCTCCGCTTAAGCTCTTCATGACAGAAAGCTTGACCCCTCTGCTTGAA	986
Qy	541	TTCTCTCAGATTCATTGTCAACAGGGCAAGAAATGATGTGAAATGACAGAAATTCAGAAC	600
Db	987	TTCTCTCAAGTCCACTGCTTACAGGGGAAAGAAATGATGTGAAATGCCAGAAATCCAGAAAT	1046
Qy	601	TTGAATGCTCTTCAACATGTTTGTGAGAAAGAGAGAGATATCTTGTAACCTCACTACTCT	660
Db	1047	GTGTTCTCTCTCCCAATGATCTCTGAGAAAGGAGAGATATCTGAACCTCGTCAAAATCC	1106
Qy	661	GAGACACCCATGTGCCGAATTCGAAACCCGCTTCCAGAAATCGATTTGGAAAGAGTTGG	720
Db	1107	GAGACCCCGTACTCCCAATTTGAGCAAAATTCAGAGATATCGGCTGAGAGGGGAGTGG	1166
Qy	721	GGTGACACCGCAAGACGCGTGTGAGATGATCCATCTCTTTTGGATCTTTGAGCA	780
Db	1167	GGTGACACGGGTGAGCGCGCTCTCGAAGATGATCCAGTCTGTGTAATCTCTTGAAGCT	1226
Qy	781	ACTGATCTTGTGACCTTGAAGAGTTCCCTGGGAAATCCCATGATGTTCAATGTTGTG	840
Db	1227	CCCGACCCGTGACTCTCGAAGAGTCTTGGATGTAGGATTTCCATATGCTTTCACATGCTGTG	1286
Qy	841	ATTCTCATCTCCCAAGATTACTTCGCTCAAGACATGTTTGGGTATCCGACACCCGT	900
Db	1287	ATCATATGCTCCCAAGATTACTTGTCTCAGAGCAAGTCTTGTGTTATCCGGAATACCGGT	1346
Qy	901	GGGCAAGTTTGTTAATCTTGGATCAAGTCCGAGCTTTGGAGATGAGATGCTCTCGCT	960
Db	1347	GGCCAGGTGTATTAATCTCGATTAAGTTGCTGCTTACGAGAGAAATCTTCAACGC	1406
Qy	961	ATTAAGCAACAGAGACTCAATCAACCCCTGAAATCCTCATTAATACAGACTTTCCT	1020
Db	1407	ATTAGCAACAGAGACTGATATTAATCTCTCGAATCTCATTAATACCTCGCTTCTTCCA	1466
Qy	1021	GATGCTGTGGAACAACATGCGGTCAACGACTTGTGAAAGTATACGAAACAGAGCACTCG	1080
Db	1467	GACGCGGTGGAAACCACTGTGTGGCAGAGGCTTGAAGAAATTTTGGGACCGAGATACCTC	1526
Qy	1081	GATATTTCTTGAAGTACCTTCAAGAACAGAAAGGAAATTTGTGAAAATGATCTCAAGA	1140
Db	1527	CACATTTCTTGCATCTCCCTTCAGAATAGAAAGGAGTGTGCTCGCAATGTGAATTTCCGG	1586
Qy	1141	TTTGAAGAAAGTGGCCAAATCTTGGAAACTTACACAGAGAGATGTTGTCAATGAATCTCC	1200
Db	1587	TTTCG---AGGTGTGCTCTATTTTGGAAAGATACAGTAGAGATGTGCGACGAACTTGTCT	1643
Qy	1261	AAAGAGTTGACCGGACCGCCAGATGTGATCTGGAAACMAACGGACGGCAATATCGTCTC	1266
Db	1644	GGAAGTTTGACGGGCAACCTGATCTGATCATCGGAACTTACAGTGAATGAAACATTTGTT	1703
Qy	1261	GCCTCTTGTCTCGCAATTAATTAAGTGTCAACATGTGCAACATGTGCCCATATGCTTTGGAG	1320
Db	1704	GCCTCTTGTGTAGCAATTAATTAAGTGTATACACAGTGTAACTAATGATGCCATATGCCCTCGAG	1763

QY	132.	AAGCAAAATATCCAGATTCCAGATATCTATTTGGAAAGCGTTGAAGCAATATCCATTCC	1380
Db	1764	AAGACGAGATACCCAGAGTCAAGACATATCTGGAGAAATTTGGAGAAAGTACCACTTCC	1823
QY	1381	TCTTCCCAATTTTCAGCTGATCTTTTGTGCAATGAAACCATACAGATTTTCATCATCCAGT	1440
Db	1824	TCTTGGCAGTTTCACTGCTGATCTTCATGCGCATGAAACGACCGGACTTCATTATCCAGCC	1883
QY	1441	ACTTTCAGAGAAATTCAGAGAGAGAGACACTGTTGGTCAATATCGAGGCCACACTGCT	1500
Db	1884	ACCTTCCAAAGAAATGCTGGAGACGAGATACAGTGGGGCAGTATGAGAGCTCACATGAC	1943
QY	1501	TTGACCTTTCCTGGTCTACCTCGGTGTGTACATGATGATGATGTTGATATCCCAATTC	1560
Db	1944	TTCACTTTCTTCGACCTTACCCAGATGTTCCAGGGATGACCGTCTTGACCCGAAATTC	2003
QY	1561	AACATTTGTTCCCTGGTGTGCTGATATGAGATATACTCTCTTACACCGAGAGAGACGG	1620
Db	2004	AACATTTGTTTCCACGAGTGTGACATGACATCTACTTGGCTTACACCGAAGAGAGCGG	2063
QY	1621	AGGTTGAAGCATTTCCATCCTGAGATGGAAGACCTTCTTTACACCAAGTTGAGAAATGAA	1680
Db	2064	CGGTTGAATATCCTTCCACCTGAGATGAGAGAACTCCTCTTGACGCGATGTTGAGAAACAG	2123
QY	1681	GAACACTTATGTGTGTCTCAATGACCGGACAAAGCCAACTTCTGTTCAATATGCCAGGCTT	1740
Db	2124	GAACACTTGTGTGTGTGAAAGATTAAGAAAGCCTATATTTATTCACATGAGGACGGCTG	2183
QY	1741	GATCGTGTCAAGAACTTAACCGGACTGTGAGTGGTGGCGGCAAGAACCCAAAGTTGCGT	1800
Db	2184	GACCGTGTCAAGAACTTGAACAGGCTTGTGAGTGGTATGAGCAAGAACTCCAAAGTTGAGG	2243
QY	1801	GAGTTGGCTTAACCTCGTAGTTGTAGTGTGATATAGCGGAAGAAATCTAAAGATTTGGAA	1860
Db	2244	GAACTGGCCAACTTGGTGTGTGTGGAGGTGACAGAGAGAAAGTATGGAAGCACTTGAA	2303
QY	1861	GAGAAAGCTGAATGAAGAAATGTTTGAGCTGATGACAAATGACAACTTGAACGGCCAA	1920
Db	2304	GAGCAAGCTGAGATGAAGAAATGTATCGACCTCATCAAAAGTACAACTGAAATGGCCAG	2363
QY	1921	TTTCAGATGAAATATCATCTCAATGAAGACAGAAATCGAAATGTTGAACCTTACCCGTAATT	1980
Db	2364	TTTCAGGTGAATTCCTCCAGATGAACCGGGTGAAGATGAGAGGCTCTACCCGTTAATC	2423
QY	1981	TGCGACACGAAAGGTGCTTTGTATACAGCTTGCAATTTGATAAAGCTTTGGATTGACAGTT	2040
Db	2424	TGTGACACGAAAGGAGTCTTCGTTCAACCGGCTATCTATGAAGCTTTGGGTTGACCGTG	2483
QY	2041	GTCGAGCAATGACTTGGCGGTTTGCCAACATTGCGCACTGTATACGGTGAACAGCCGAG	2100
Db	2484	GTTGAGGCGCAATGACTTGTGATTTGCCAACCTTTGCCACTTGCAATGGTGAACGACGTAG	2543
QY	2101	ATTATTTGTCATGGGAAATCTGTTTCCAAATTTGATTCCTTACCATGGTGATTCAAAGTCT	2160
Db	2544	ATTCATTTGTGATGCAAAATCGGCGTACCAATTTGATCTTACCAATGGTGAACAGGCGGCC	2603
QY	2161	GACATATCTCGTCAATTTCTTTGAAAGTGTAAAGAAAGATCCATCTCACTGGGATTAAGATC	2220
Db	2604	GAGCTTTTGTGAATCTTCTTCAACAAAGTGCAGATTTGACCAAGTCCCACTGGAGCAGATC	2663
QY	2221	TCCCAAGAGGCTTGAAGCAAAATAGAGAGAGATATCATGAGAAATTTTACTCGGAGAGA	2280
Db	2664	TCAAGAGGTGCATGACGAGAAATTTGAAGAGAAATATCATGAGAAATATATTTGAGAGG	2723
QY	2281	CTATTTGACCTTGACAGAGATGTATGATTTTGGAGAGATGTTTCCAACTTTGAACCGCGT	2340
Db	2724	CTGTTGAACCTGACCTGCGGTGTATGGCTTTTGGAGAGATGTGACTAACTTTGATCGGCGC	2783
QY	2341	GAGAGTGTGCTTACCTTGAGATGTTTATGCTCTTAAATGACGTAAGCTGAGGCTGAATCA	2400
Db	2784	GAGAGTGTGCGGATACCTTGAAGATGTTCTATGCTCTCAAGTATGCCCACTGGCACAGCT	2843
QY	2401	GTTCCATTTGGCAGAGAGTAAA 2422	

Wed Jun 9 11:54:30 2004

us-10-003-405-1.rge

Page 25

Db 2844 GTCTCCGCGCTGTCGAGTAA 2865

Search completed: June 5, 2004, 05:44:35
Job time : 16241 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - pteotein search, using sw model

Run on: June 2, 2004, 14:49:54 ; Search time 22 Seconds
(without alignments)
1891.388 Million cell updates/sec

Title: US-10-003-405-2

Perfect score: 4247

Sequence: 1 MAERLTVHSIRERLDETL.....EMFVLRKYLAEVPLAE 806

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3643.5	85.8	805	US-09-598-401C-77	Sequence 77, Appl
2	2841	66.9	766	US-08-553-436A-8	Sequence 8, Appl
3	1747	41.1	806	US-08-684-005-2	Sequence 2, Appl
4	506	11.9	125	US-09-598-401C-76	Sequence 76, Appl
5	494.5	11.6	720	US-09-394-272-14	Sequence 14, Appl
6	437	10.3	1059	US-09-394-272-5	Sequence 5, Appl
7	436	10.3	1068	US-08-429-054A-11	Sequence 11, Appl
8	436	10.3	1068	US-08-718-777-7	Sequence 7, Appl
9	436	10.3	1068	US-09-051-341-7	Sequence 7, Appl
10	435	10.3	1083	US-09-394-272-8	Sequence 8, Appl
11	435	10.3	1083	US-09-394-272-11	Sequence 11, Appl
12	431	10.1	1056	US-09-394-272-1	Sequence 1, Appl
13	421.5	9.9	1045	US-08-553-436A-6	Sequence 6, Appl
14	421.5	9.9	1045	US-09-394-272-7	Sequence 7, Appl
15	419	9.9	1054	US-08-356-354-4	Sequence 4, Appl
16	419	9.9	1054	US-08-778-656-4	Sequence 4, Appl
17	418.5	9.9	1081	US-09-394-272-4	Sequence 4, Appl
18	413.5	9.7	1053	US-09-394-272-6	Sequence 6, Appl
19	408	9.6	1059	US-09-697-367-16	Sequence 16, Appl
20	407.5	9.6	963	US-09-394-272-12	Sequence 12, Appl
21	407.5	9.6	963	US-09-394-272-13	Sequence 13, Appl
22	407	9.6	1057	US-09-697-367-23	Sequence 23, Appl
23	407	9.6	1057	US-09-394-272-2	Sequence 2, Appl
24	406	9.6	1057	US-08-853-948B-2	Sequence 2, Appl
25	400	9.4	1084	US-09-394-272-9	Sequence 9, Appl
26	394	9.3	1054	US-09-394-272-3	Sequence 3, Appl
27	387.5	9.1	1049	US-09-394-272-10	Sequence 10, Appl

28	367.5	8.7	846	1	US-08-356-354-2	Sequence 2, Appl
29	367.5	8.7	846	2	US-08-778-656-2	Sequence 2, Appl
30	305	7.2	908	1	US-08-356-354-6	Sequence 6, Appl
31	305	7.2	908	2	US-08-778-656-6	Sequence 6, Appl
32	286.5	6.7	668	4	US-09-697-367-2	Sequence 2, Appl
33	222.5	5.2	365	4	US-09-697-367-24	Sequence 24, Appl
34	209.5	4.9	365	4	US-09-697-367-8	Sequence 8, Appl
35	192.5	4.5	341	3	US-08-853-948B-4	Sequence 4, Appl
36	189.5	4.5	343	3	US-08-853-948B-1	Sequence 3, Appl
37	184.5	4.3	343	3	US-08-853-948B-10	Sequence 10, Appl
38	182.5	4.3	358	4	US-09-697-367-20	Sequence 20, Appl
39	182	4.3	348	3	US-08-853-948B-5	Sequence 5, Appl
40	153.5	3.6	502	4	US-09-134-001C-4511	Sequence 4511, Ap
41	153	3.6	442	4	US-09-489-039A-8158	Sequence 8158, Ap
42	150.5	3.5	418	4	US-09-107-532A-6073	Sequence 6073, Ap
43	150	3.5	393	4	US-09-252-991A-19043	Sequence 19043, A
44	148.5	3.5	210	4	US-09-697-367-10	Sequence 10, Appl
45	142.5	3.4	369	4	US-09-543-681A-7514	Sequence 7514, Ap

ALIGNMENTS

```
RESULT 1
US-09-598-401C-77
; Sequence 77, Application US/09598401C
; Patent No. 6596925
; GENERAL INFORMATION:
; APPLICANT: Perera, J. Ranjan
; APPLICANT: Sagleton, Clare
; APPLICANT: Rice, Stephen J.
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1036C2
; CURRENT FILING DATE: 2000-06-20
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: U.S. No. 6596925 60/146,591
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: PCT/NZ00/00018
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: U.S. No. 6596925 09/276,599
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 805
; ORGANISM: Eucalyptus grandis
US-09-598-401C-77
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
Query Match      85.8%; Score 3643.5; DB 4; Length 805;
Best Local Similarity 85.0%; Pred. No. 0;
Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;

QY 1 MAERLTVHSIRERLDETLAHRREILALSRIGKKGILQHOQITLPEPAIPENRK 60
DB 1 MADRLTSHSIRERLDETLAHRREILALSRIGKKGILQHOQITLPEPAIPENRK 60
QY 61 KLANGAFPEVLKASOEAVLPPWVALAVRPPGWVEYIRVWVHALVVELTVAEYLAFKE 120
DB 61 KLNGAFPEVLKASOEAVLPPWVALAVRPPGWVEYIRVWVHALVVELTVAEYLAFKE 120
QY 121 EIVDSSNGNRYLEIDPEPNSSPPRLTSSIGNGVEFLNRHLSAKLPHDKESHAPLE 180
DB 121 EIVDSSNGNRYLEIDPEPNSSPPRLTSSIGNGVEFLNRHLSAKLPHDKESHAPLE 180
QY 121 EIVDSSNGNRYLEIDPEPNSSPPRLTSSIGNGVEFLNRHLSAKLPHDKESHAPLE 180
DB 121 EIVDSSNGNRYLEIDPEPNSSPPRLTSSIGNGVEFLNRHLSAKLPHDKESHAPLE 180
QY 181 FLRVCHGKMMNDRIQNALQHVLRKBEYVGTLPPTPCAFEFHROEIGLEBGW 240
DB 181 FLRVCHGKMMNDRIQNALQHVLRKBEYVGTLPPTPCAFEFHROEIGLEBGW 240
QY 241 GDTARVLEMIQLLIDLLEATDPCTLEKFLRIPWVFNVVILTPHGYFAQDNVIGYPTDG 300
DB 241 GDTARVLEMIQLLIDLLEATDPCTLEKFLRIPWVFNVVILTPHGYFAQDNVIGYPTDG 300
```

```

QY 301 GQVYIILDOVALENEMILRIKQOGLNTPRILITRLLPDAVGTTCCGRLEKVTGHS 360
D 301 GQVYIILDOVALENEMILRIKQOGLNTPRILITRLLPDAVGTTCCGRLEKVTGHS 360
QY 361 DILAVPTEKGIYKRWISREPKWPLYETTEDVAHISKEIAGTDPDIIIGNSDGNIV 420
D 361 DILAVPTEKGIYKRWISREPKWPLYETTEDVAHISKEIAGTDPDIIIGNSDGNIV 420
QY 421 ASLHAKRGVTOCTTAHLEKTKYPSDIYWKLEDKYHFGSCQFADLIFANNHTDFTTS 480
D 421 ASLHAKRGVTOCTTAHLEKTKYPSDIYWKLEDKYHFGSCQFADLIFANNHTDFTTS 480
QY 481 TFOEIASGKDTYGVQYESHATFLLPGLYRVVHGIDVDFDKFNIVSPGADMEIYFPYTEER 540
D 481 TFOEIASGKDTYGVQYESHATFLLPGLYRVVHGIDVDFDKFNIVSPGADMEIYFPYTEER 540
QY 541 RLKHFHPEIEDLLTYKVENEBHLCVLDRNKPILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
D 541 RLKHFHPEIEDLLTYKVENEBHLCVLDRNKPILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
QY 601 ELANLVVVGDRRKESKDLSEKEMKMFELIDKNLNGOFPMISSQNNRIRNVELYRYI 660
D 601 ELANLVVVGDRRKESKDLSEKEMKMFELIDKNLNGOFPMISSQNNRIRNVELYRYI 660
QY 661 CDTKGFVQPALYEAFGLTVEAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQA 720
D 661 CDTKGFVQPALYEAFGLTVEAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQA 720
QY 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKIYSERLTLTGYYGFWKHSNLER 780
D 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKIYSERLTLTGYYGFWKHSNLER 780
QY 781 ESRRLYEMFYALKYKLAESVPLAE 806
D 781 ESRRLYEMFYALKYKLAESVPLAE 806

```

RESULT 2

US-08-553-436A-8
Sequence 8, Application US/08553436A

Patent No. 5866790

GENERAL INFORMATION:

APPLICANT: HESSE, Holger

APPLICANT: MULLER-ROBER, Bernd

TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE

TITLE OF INVENTION: PREPARATION OF SUGAR BEET WITH CHANGED SUCROSE

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ostrolenk, Faber, Garb & Soffen

STREET: 1180 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: US

ZIP: 10036-8403

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/553,436A

FILING DATE: 17-NOV-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP94/01671

FILING DATE: 20-MAY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P 4317596.1

FILING DATE: 24-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Weilman, Edward

```

/ REGISTRATION NUMBER: 24,735
/ REFERENCE/DOCKET NUMBER: P/951-117
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 382-0700
/ TELEFAX: (212) 382-0888
/ TELEX: 236925
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 766 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-553-436A-8

```

Query Match 66.9%; Score 2841; DB 2; Length 766;
Best Local Similarity 72.4%; Pred. No. 6.5e-269;
Matches 541; Conservative 90; Mismatches 112; Indels 4; Gaps 4;

```

QY 62 LANGAFPEVILKASGEAIVLPPWVALAVRPRPGWEYIRVNVALVEELTYAEYLHFEKE 121
D 62 LANGAFPEVILKASGEAIVLPPWVALAVRPRPGWEYIRVNVALVEELTYAEYLHFEKE 121
QY 7 LSDGPFSEVLRSAGEAIVPPFVALAVRPRPGWEYIRVNVALVEELTYAEYLHFEKE 66
D 7 LSDGPFSEVLRSAGEAIVPPFVALAVRPRPGWEYIRVNVALVEELTYAEYLHFEKE 66
QY 122 LVDSGSGNFVLELDEPFPENSSFPRLTSKISGVGEFELNRLSAKLPHDKESMPLLEF 181
D 122 LVDSGSGNFVLELDEPFPENSSFPRLTSKISGVGEFELNRLSAKLPHDKESMPLLEF 181
QY 67 LVDSGADHDVYVLELDEPFPENSSFPRLTSKISGVGEFELNRLSAKLPHDKESMPLLEF 126
D 67 LVDSGADHDVYVLELDEPFPENSSFPRLTSKISGVGEFELNRLSAKLPHDKESMPLLEF 126
QY 182 LRVHCHGKMMMLNDRIQNLALQHLRKAEEYLTLPPTPCAETFEHFOEIGLERGNG 241
D 182 LRVHCHGKMMMLNDRIQNLALQHLRKAEEYLTLPPTPCAETFEHFOEIGLERGNG 241
QY 127 LRVHCHGKMMMLNDRIQNLALQHLRKAEEYLTLPPTPCAETFEHFOEIGLERGNG 186
D 127 LRVHCHGKMMMLNDRIQNLALQHLRKAEEYLTLPPTPCAETFEHFOEIGLERGNG 186
QY 242 DTAERVLEMTQLDLLEAIDPCTLEKFLGRIPMVFNVVITTPRGYPAQDNVLCYPTGG 301
D 242 DTAERVLEMTQLDLLEAIDPCTLEKFLGRIPMVFNVVITTPRGYPAQDNVLCYPTGG 301
QY 187 DTAERVLEMTQLDLLEAIDPCTLEKFLGRIPMVFNVVITTPRGYPAQDNVLCYPTGG 246
D 187 DTAERVLEMTQLDLLEAIDPCTLEKFLGRIPMVFNVVITTPRGYPAQDNVLCYPTGG 246
QY 302 GQVYIILDOVALENEMILRIKQOGLNTPRILITRLLPDAVGTTCCGRLEKVTGHS 361
D 302 GQVYIILDOVALENEMILRIKQOGLNTPRILITRLLPDAVGTTCCGRLEKVTGHS 361
QY 247 QIVYIILDOVALENEMILRIKQOGLNTPRILITRLLPDAVGTTCCGRLEKVTGHS 306
D 247 QIVYIILDOVALENEMILRIKQOGLNTPRILITRLLPDAVGTTCCGRLEKVTGHS 306
QY 362 ILRVPRTEKGIYKRWISREPKWPLYETTEDVAHISKEIAGTDPDIIIGNSDGNIV 421
D 362 ILRVPRTEKGIYKRWISREPKWPLYETTEDVAHISKEIAGTDPDIIIGNSDGNIV 421
QY 307 ILRVPRTEKGIYKRWISREPKWPLYETTEDVAHISKEIAGTDPDIIIGNSDGNIV 365
D 307 ILRVPRTEKGIYKRWISREPKWPLYETTEDVAHISKEIAGTDPDIIIGNSDGNIV 365
QY 422 SLAHKRGVTOCTTAHLEKTKYPSDIYWKLEDKYHFGSCQFADLIFANNHTDFTTS 481
D 422 SLAHKRGVTOCTTAHLEKTKYPSDIYWKLEDKYHFGSCQFADLIFANNHTDFTTS 481
QY 366 SLHSHKRGVTOCTTAHLEKTKYPSDIYWKLEDKYHFGSCQFADLIFANNHTDFTTS 425
D 366 SLHSHKRGVTOCTTAHLEKTKYPSDIYWKLEDKYHFGSCQFADLIFANNHTDFTTS 425
QY 482 TFOEIASGKDTYGVQYESHATFLLPGLYRVVHGIDVDFDKFNIVSPGADMEIYFPYTEER 540
D 482 TFOEIASGKDTYGVQYESHATFLLPGLYRVVHGIDVDFDKFNIVSPGADMEIYFPYTEER 540
QY 426 YOEIAGTKNTVGOYESHATFLLPGLYRVVHGIDVDFDKFNIVSPGADMEIYFPYTEER 485
D 426 YOEIAGTKNTVGOYESHATFLLPGLYRVVHGIDVDFDKFNIVSPGADMEIYFPYTEER 485
QY 541 RLKHFHPEIEDLLTYKVENEBHLCVLDRNKPILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
D 541 RLKHFHPEIEDLLTYKVENEBHLCVLDRNKPILFTMPRLDRVKNLTGLVEMCGKNPKLR 600
QY 486 CLTSLHKLIDGLLKPENNEHIGVLDTSKPIIFSMARLDLRVKNLTGLVEMCGKNPKLR 545
D 486 CLTSLHKLIDGLLKPENNEHIGVLDTSKPIIFSMARLDLRVKNLTGLVEMCGKNPKLR 545
QY 601 ELANLVVVG-DRRKESKDLSEKEMKMFELIDKNLNGOFPMISSQNNRIRNVELYRYI 659
D 601 ELANLVVVG-DRRKESKDLSEKEMKMFELIDKNLNGOFPMISSQNNRIRNVELYRYI 659
QY 546 ELANLVVVG-DRRKESKDLSEKEMKMFELIDKNLNGOFPMISSQNNRIRNVELYRYI 605
D 546 ELANLVVVG-DRRKESKDLSEKEMKMFELIDKNLNGOFPMISSQNNRIRNVELYRYI 605
QY 660 ICDTKGAFVQPALYEAFGLTVEAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQA 719
D 660 ICDTKGAFVQPALYEAFGLTVEAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQA 719
QY 606 ICDTKGAFVQPALYEAFGLTVEAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQA 664
D 606 ICDTKGAFVQPALYEAFGLTVEAMTCGLPTFATCNGCPAEIIVHSGSGFNIDPYHGDQA 664
QY 720 ADILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKIYSERLTLTGYYGFWKHSNLER 779
D 720 ADILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKIYSERLTLTGYYGFWKHSNLER 779
QY 665 AEKMTPEFVKCRDNDPNWTIKSAGGLRIKRIYRWKYSERLMTLAGYGFWMKYSKLER 724
D 665 AEKMTPEFVKCRDNDPNWTIKSAGGLRIKRIYRWKYSERLMTLAGYGFWMKYSKLER 724
QY 780 ESRRLYEMFYALKYKLAESVPLAE 806
D 780 ESRRLYEMFYALKYKLAESVPLAE 806
QY 725 RETRRLYEMFYALKYKLAESVPLAE 751
D 725 RETRRLYEMFYALKYKLAESVPLAE 751

```

RESULT 3

US-08-684-005-2

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/429,054A
FILING DATE: 26-APR-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 842,337
FILING DATE: 20-March-1992
APPLICATION NUMBER: PCT/FR 91/00593
FILING DATE: 18-July-1991
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: French 90402094.9
FILING DATE: 20-July-1990
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Charles A. Muellerian
REGISTRATION NUMBER: 19,683
REFERENCE/DOCKET NUMBER: 146.1137
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 661-8000
TELEFAX: (212) 661-8002
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1068
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Unknown
MOLECULE TYPE: Peptide
US-08-429-054A-11

Query Match 10.3%; Score 436; DB 2; Length 1068;
Best Local Similarity 25.2%; Pred. No. 4.5e-33;
Matches 167; Conservative 114; Mismatches 213; Indels 170; Gaps 29;
Qy 185 HCHKKMMMLNDRIQNLNA-----LQHVLRKAEVY-----GTLPP-ETPCAE 226
Db 97 HLAKKKKQLELEIGRISARRKEQVRRATEDLAEDLSGEKGDITIGELAPVETTKK 156
Qy 227 FEHRPOEIGLERGMDTAERVLEMIQLLDLLEATDPCTLEKFLGRIPMVFNVILTPHG 286
Db 157 FQNRSDTLV---WSDDNK-----EKKL-----YIVLISVHG 185
Qy 287 YFAQDNV-LGV-PDTGGGVVYILDQVRALENEMMLRIKQGLNITP---RILITRL-- 339
Db 186 LVRGEMELGRSDYTGQGVKYVELARA-----MSMMPGVYRVDLFTROYSS 232
Qy 340 PD-----AVGTTCCGRLKGVYGTESDILVVPRTKEGIYKRWISREKWPY 387
Db 233 PDVDMSYEPTPEMLCAGSNDGEGESG---AYIVRIPCGRDKYLKK---BALMPY 284
Qy 388 LETYTD-VAH--EISKEL-----HGTPLD---IIGNSDGNIVASLAHKLGVQCTIA 436
Db 285 LQEPVDAALAHILNLSKALGEOVNGRPVLPVIGHYADAGVALLSGALNVPMVYLG 344
Qy 437 HALEKTK---YPSDIIYWKKLEDKYIHNSCOFTADLPMMNHTDPIITSTFOEIASGKTV 492
Db 345 HSLGRNKEQLLKQGRMSKEEIDSTYKIMRIEGELALDASELIVTSTROI---DEQW 401
Qy 493 GQYESHTAFTLPGYLRVYHGDV-----FDDKFNIVSGADMEIYF 533
Db 402 GLYD-----GFDVKLEKVLRAARRGVSCHGKRYMRVNVIPGMDPSNVV 446
Qy 534 PY-----TEEKRLKHFPHEIEDLLYTKVNEEHLCVLNDNRKPIFLTPMLIDRVK 584
Db 447 VHEIDIDGQDVKDDIVGLEGASPKSMPIWAEV-----MRFLTNHKKMILALSPDPKK 501
Qy 585 NLTGIVEMCGKPKRELANLVVGGDRKESKDLSEKAE-----MKGMFELIDKYNLNG 639
Db 502 NITTLVKAFGECRPLRELANLTLLIMGNR---DDIDMSAGNASVLTTLKLIKIDKDYDG 557
Qy 640 QPRWISQMNRIIRNVELRYICDTGAFVQPALYEAFGLTVEANTGCLPFPATNGGPA 699

Db 558 SVAF-PKHNOADVPEIRLAAKKGVFINVALVEPFLTILEAAAGLPIVATNGGPV 616
Qy 700 EIVHSGSGFNIDPYHGDQADILVDPEKCKPSHMDKISOGGLKRIEKTMYKISE 759
Db 617 DITNALNNGLLVDPHDQALADALL-----KLVAQKNLMQECRRNGLRNI-HLYSWPEHCR 671
Qy 760 RLIT 763
Db 672 TYLT 675

RESULT 8
US-08-718-777-7
Sequence 7, Application US/08718777
Patent No. 5981852

GENERAL INFORMATION:
APPLICANT: Van Asche, C.
APPLICANT: Lando, D.
APPLICANT: Bruneau, J. M.
APPLICANT: Voelker, T.
APPLICANT: Gervais, M.
TITLE OF INVENTION: MODIFICATION OF SUCROSE
TITLE OF INVENTION: PHOSPHATE
TITLE OF INVENTION: SYNTHASE IN PLANTS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Barbara Rae-Venter
STREET: 260 Sheridan Avenue, Suite 440
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718,777
FILING DATE: NOT YET ASSIGNED
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,471
FILING DATE: 27-DEC-1993

ATTORNEY/AGENT INFORMATION:
NAME: Barbara Rae-Venter
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: CGNE.072.02US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)328-4400
TELEFAX: (415)328-4477

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1068 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-718-777-7

Query Match 10.3%; Score 436; DB 2; Length 1068;
Best Local Similarity 25.2%; Pred. No. 4.5e-33;
Matches 167; Conservative 114; Mismatches 213; Indels 170; Gaps 29;

Qy 185 HCHKKMMMLNDRIQNLNA-----LQHVLRKAEVY-----GTLPP-ETPCAE 226
Db 97 HLAKKKKQLELEIGRISARRKEQVRRATEDLAEDLSGEKGDITIGELAPVETTKK 156
Qy 227 FEHRPOEIGLERGMDTAERVLEMIQLLDLLEATDPCTLEKFLGRIPMVFNVILTPHG 286
Db 157 FQNRSDTLV---WSDDNK-----EKKL-----YIVLISVHG 185
Qy 287 YFAQDNV-LGV-PDTGGGVVYILDQVRALENEMMLRIKQGLNITP---RILITRL-- 339

```

Db 186 LVGRNEMELGRDSTGGQVKKVVELARA-----MSMMPGVYRDLFTFRQVSS 232
Qy 340 PD-----AVGTTCGQRLEKVVGTESHDILRVFPRTKEKIVKRWISREKVPY 387
Db 233 PDVMSYGEPTMLCAGSNDGEGMSGG---AYIVRIPCGPRDKYLK-----EALMPY 284
Qy 388 LETYED-VAH--EISKEL-----HGTPLD---ITGNXSDGNIYASLLAHKLGVTQCTIA 436
Db 285 LQEFVDGALAHILNMSKALGEOVNGRPVLPYVIGHYADAGDVALLSGALNVPMLVG 344
Qy 437 HALEKTK-----YPSDDIYWKLEBDKYNHSCFTADLPAMNHTDFTITSTFOEIASKDTV 492
Db 345 HSLGNKKEQLKQGRMSKEEIDSTYKIMRIEGBELADLASELVITSTROEI---DEQW 401
Qy 493 GQYESHTAFTLPGLVRVVHGIDV-----FDPKNIVSPGADMEIYF 533
Db 402 GLYD-----GPDVLEKYLBARARGVSGHGRYMRMVVITPPGMDFSNVV 446
Qy 534 PY-----TEKRRLKHPHEIEDLLYTKVNEBHLCVLDRNKPILFTMRDLRVK 584
Db 447 VHEDIDGDGVKDDIVGLEGASPKSMPPIMAEV-----MRFLTNPHKMILALSRPDPK 501
Qy 585 NITGLVEMCGKNPRLRELANLVVGGDRKESKDLSEKAB-----MKMFELIDKYNLNG 639
Db 502 NITTLVKAFGECRPLRELANLTILMGNR-----DDIDMSAGNASVLTTLVKLIDKIDYLG 557
Qy 640 QFRWISSQMRIRNVELRYICDTKGAFVOPALYEAFLVVEAMTCGLPTFATCNGGPA 699
Db 558 SVAF-PKHNGADVBEIRLAARKKGVFINPALVPEFGLTLEBAAGLPIVATKNGSPV 616
Qy 700 EIIYHSGSGFNIDPYHGDQADILVDFEKCCKDPSHMDKISQGLKRIEKKYTWKIYSE 759
Db 617 DITNALNNGLLVDPHDQNAIDALL-----KLVAADKNLMQECRRNGLRNI-HLYSPBCHR 671
Qy 760 RLTT 763
Db 672 TYLT 675

RESULT 9
US-09-051-341-7
; Sequence 7, Application US/09051341
; Patent No. 6124528
; GENERAL INFORMATION:
; APPLICANT: Shemmaker, C. K.
; TITLE OF INVENTION: MODIFICATION OF SOLUBLE SOLIDS USING
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE ENCODING SEQUENCE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rae-Venter Law Group, P.C.
; STREET: 260 Sheridan Avenue, Suite 440
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/051,341
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/17351
; FILING DATE: 25-OCT-1996
; APPLICATION NUMBER: US 08/549,016
; FILING DATE: 27-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/372,200
; FILING DATE: 12-JAN-1995

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter, Ph.D.,
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: CGNE.110.02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)328-4400
; TELEFAX: (415)328-4477
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1068 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-051-341-7

Query March 10.3%; Score 436; DB 3; Length 1068;
Best Local Similarity 25.2%; Pred. No. 4,56-33;
Matches 167; Conservative 114; Mismatches 213; Indels 170; Gaps 29;

Qy 185 HCHGKNMMLNDRIQNLNA-----LQHVLRKAEEYL-----GTLPP-ETPCAE 226
Db 97 HLAKKKQLELEGIQISARKKEQVREATEDLADLSBGEKGTIGELAPVETTKK 156
Qy 227 FEHFQETIGERCGDTAERVLMTQLLDLEATDECTLEKPLGRTPMPVNVYILPHG 286
Db 157 FQRFNSDLTV-----WSDNKK-----EKKL-----YVLIVSHG 185
Qy 287 YFAADNV-LGY-PPYGGQVYIILDOVRALENEMLRIKOQGLNITP---RLITITRL-- 339
Db 186 LVGRNEMELGRDSTGGQVKKVVELARA-----MSMMPGVYRDLFTFRQVSS 232
Qy 340 PD-----AVGTTCGQRLEKVVGTESHDILRVFPRTKEKIVKRWISREKVPY 387
Db 233 PDVMSYGEPTMLCAGSNDGEGMSGG---AYIVRIPCGPRDKYLK-----EALMPY 284
Qy 388 LETYED-VAH--EISKEL-----HGTPLD---ITGNXSDGNIYASLLAHKLGVTQCTIA 436
Db 285 LQEFVDGALAHILNMSKALGEOVNGRPVLPYVIGHYADAGDVALLSGALNVPMLVG 344
Qy 437 HALEKTK-----YPSDDIYWKLEBDKYNHSCFTADLPAMNHTDFTITSTFOEIASKDTV 492
Db 345 HSLGNKKEQLKQGRMSKEEIDSTYKIMRIEGBELADLASELVITSTROEI---DEQW 401
Qy 493 GQYESHTAFTLPGLVRVVHGIDV-----FDPKNIVSPGADMEIYF 533
Db 402 GLYD-----GPDVLEKYLBARARGVSGHGRYMRMVVITPPGMDFSNVV 446
Qy 534 PY-----TEKRRLKHPHEIEDLLYTKVNEBHLCVLDRNKPILFTMRDLRVK 584
Db 447 VHEDIDGDGVKDDIVGLEGASPKSMPPIMAEV-----MRFLTNPHKMILALSRPDPK 501
Qy 585 NITGLVEMCGKNPRLRELANLVVGGDRKESKDLSEKAB-----MKMFELIDKYNLNG 639
Db 502 NITTLVKAFGECRPLRELANLTILMGNR-----DDIDMSAGNASVLTTLVKLIDKIDYLG 557
Qy 640 QFRWISSQMRIRNVELRYICDTKGAFVOPALYEAFLVVEAMTCGLPTFATCNGGPA 699
Db 558 SVAF-PKHNGADVBEIRLAARKKGVFINPALVPEFGLTLEBAAGLPIVATKNGSPV 616
Qy 700 EIIYHSGSGFNIDPYHGDQADILVDFEKCCKDPSHMDKISQGLKRIEKKYTWKIYSE 759
Db 617 DITNALNNGLLVDPHDQNAIDALL-----KLVAADKNLMQECRRNGLRNI-HLYSPBCHR 671
Qy 760 RLTT 763
Db 672 TYLT 675

RESULT 10
US-09-394-272-8
; Sequence 8, Application US/09394272
; Patent No. 6472588
; GENERAL INFORMATION:

```


APPLICANT: Haigler, Candace H.
APPLICANT: Holaday, A. Scott
TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
TITLE OF INVENTION: EXPRESSION OF SUCROSE PHOSPHATE SYNTHASE
FILE REFERENCE: 201304/1000
CURRENT APPLICATION NUMBER: US/09/394,272
CURRENT FILING DATE: 1999-09-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 1068
TYPE: PRT
ORGANISM: Zea mays
US-09-394-272-8

Query Match 10.3%; Score 436; DB 4; Length 1068;
Best Local Similarity 25.2%; Pred. No. 4.5e-33;
Matches 167; Conservative 114; Mismatches 213; Indels 170; Gaps 29;
QY 185 HCHKKNMMLNDRIQNLNLA-----LOHVLARKAEYL-----GTLPP-ETPCAE 226
DB 97 HLAKKKOLELEGIQRIKARKKEQOVREATEDLAEDISEGKDDTIGELAPVEITKKK 156
QY 227 FEHFOEIGLERMGDPTAERVLEMIQLDLLEATDPTCTLEKFLGRIPWFVNVVILTPHG 286
DB 157 FGRNFSDLTV---WSDDNK-----EKKL-----YIVLISVHG 185
QY 287 YFAQONV-LGY-PTDGGOVVYLLDOVRALENMLRIKQGLNITP---RIIITRL-- 339
DB 186 LVRGEMELGRDSDTGGQVYVELARA-----MSMVGVRVLDLFROYSS 232
QY 340 PD-----AVGTTCGGRLEKVGTEHSDILRVPRTEKGIYKRWISREKWPY 387
DB 233 PDVMSYGEPTMLCAGSNDGEGMGSGS---AYIVRIPCGRDYLLK---BALWY 284
QY 388 LETYED-VAH--EISKEP-----HGTPLD---IIGNSDGNIVASLAHKLGVTOCTIA 436
DB 285 LQEPFDGALAHILNLSKALGEOVNGRPVLPVIGHVADAGDAVALLSGALNVMLVLTG 344
QY 437 HALEKTK---YPSDIYWKLEDKYHSCQFTADLFAMNHTDPIITSTFOEIAISKDTV 492
DB 345 HSLGRNKEQLKQGRMSKEEIDSTYKIMRIEGEELADASELVITSTROBI---DEQM 401
QY 433 GQYEHNTAFTLBGLRVVHGIDV-----FDDKENIVSGADMEIYF 533
DB 402 GLYD-----GFDVKLEKVLRAPARRGVSGHGRYMPRVVITPGMDFSNV 446
QY 534 PY-----TEKRLKHFPEIEDLLTYTVENEHLCVANDRNKPIILFTMPRLDRVK 584
DB 447 VHEIDIDGDDVVDIVGLEGASPKSMPTWAEV---MRPLTNHKKVILALSRDPKX 501
QY 585 NLGLVEMCGKNPKLRELANLVVVGDRRKESKDLSEKAE---MKKFBELIDRYNLNG 639
DB 502 NITTLVKAFGECRPLRELANLTLLINGNR---DDIDDSAGNASVLTTLVKLIDKYDLYG 557
QY 640 QFRWISSQMRIRANVELRYICDTGCAFYQAPALYEAFGITVYEANTCGIPATNCGSPA 699
DB 588 SVAF-PKHNOADVEIYRLAAMKGVFNPLVPEFGITLLEAAHAGPIVATNCGGV 616
QY 700 EIIIVGSGFNIDPYHQADILVDFEKKCKDDSHMDKISQGLKRIEKKYTKWISYE 759
DB 617 DITNALNGLVDPHQDAIADALL---KLVAADKNLMQECRRNGLRNI-HLISVPEHCR 671
QY 760 RLTT 763
DB 672 TYLT 675

RESULT 11
US-09-394-272-11
Sequence 11, Application US/09394272
Patent No. 6472588
GENERAL INFORMATION:

APPLICANT: Haigler, Candace H.
APPLICANT: Holaday, A. Scott
TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
TITLE OF INVENTION: EXPRESSION OF SUCROSE PHOSPHATE SYNTHASE
FILE REFERENCE: 201304/1000
CURRENT APPLICATION NUMBER: US/09/394,272
CURRENT FILING DATE: 1999-09-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 1083
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-394-272-11

Query Match 10.3%; Score 435.5; DB 4; Length 1083;
Best Local Similarity 24.2%; Pred. No. 5.1e-33;
Matches 188; Conservative 128; Mismatches 272; Indels 189; Gaps 36;
QY 120 BELVDSGNGNPNVLELDEPFN-----SSFPRTLSK-----SIGNVGFILNR--HL 164
DB 44 QEKVFGDMNGDHOEKVSPIKFYEEVAVNSFSDLYTKIMIKYATNTERSRNLENI 103
QY 165 SAKLFHDKESMHP-----LEPLRVCHKKNMMLNDRIQNLNALQHLRKAE---E 213
DB 104 CMRIWHLARKKKQIYWDGVRLSKRRIEREGRDAEDLSESEKGDKNDEKESX 163
QY 214 YLGLT-PPETPCAEFHRFQSIGLRGKG--DTRERVLEMIQLDLLEATDPTCTLEKFL 270
DB 164 VTTLEPPRDHMPRIRSEMOI-----WSEDDKSRLNYILIR-----QVEIGF 207
QY 271 GRIPVFNVVI-----LTP-----HGYPADONV-LGY-PTDGGOVVYLLDOVRAL 314
DB 208 SDFVFMVNLVGLTWCLYLPFCFTNCSHGLVGRGEMELGDSDTGGQVYVELARALA 267
QY 315 NEMLRIRKQGLNITPRLIITRLPD-AVGTTCGGRLEKV-YGTEHSD---ILRVPF 367
DB 268 N-----TEGVH---RVDLLTROISPEVDYSGEPVEMLSCEPESGSGSYIIRIPC 317
QY 368 RTEKGIYKRWISREKWPYLETYTE-----DVAEHISKELHG---TPDLIIGNKXD 416
DB 318 ---GSRDKYIPK-ESLMPHPIPEFVDGALNHIIVSIRASIGOVNGKPIWPIVIGHYAD 372
QY 417 GNIVASLALHKLGYQCTIAHALEKTKY---PBSDIYWKLEDKYHSCQFTADLFAMN 472
DB 373 AGEVVAHLAALNPMVLTGSLGRNKEQLQCGRITREDIDRTYKIMRRIEAEQSLD 432
QY 473 HTDFIITSTFOEIAISKDTVQYESHNTAFTLBGLRVVHGIDV----- 515
DB 433 AAEMVVTSTROEID---AQW-----GLY---DGRDILKELRLVRRRGVSGCL 474
QY 516 --FDDKENIVSGADMEIYFPT-----BEKRLKHFPEIEDLLYTK 556
DB 475 GRYPWRVNVIPPGMD---FSYVLTQDSQEPDGLKSLIGEDRNDIKKPVPIW----- 524
QY 557 VENENHLCVANDRNKPIILFTMPRLDRVKNLGLVEMCGKNPKLRELANLVVVGDRRKES 616
DB 525 ---SEIMPFENPHKPTLLASRPDKKNVTTLVAFECOPPLRELANLVILGNR--- 577
QY 617 KDLE-----FAEMKMFELIDRYNLNGQFWMISSQMRIRANVELRYICTGKAFVOPA 671
DB 578 DDIEMRSSSVLMNVKLIDQYDLYGQVAV-PKHNGQSEVPDLYRIALATKGVFINPA 636
QY 672 LYEAFTLVBEAMTCGLPTPATNCGSPAEIIVHSGSGFNIDPYHQDQADILVDFEKKCK 731
DB 637 LVPEPGLTLEAALVGLPIVATNCGPVDIVKALNGLVDPHQDAISDALL---KLVA 692
QY 732 KDSHMDKISQGLKRIEKKYTKWISERLLTLTGCVYFMWGVNSLBERESRRLYEM 788
DB 693 ANKHLMAECRKNGLKNI-HRFSWPEHCRNYLS-----HVECHRNHPTSSLDI 739

RESULT 12

Db 398 D-PVLERKLRAMKKGVCSCYGRFMRPMVYIPRGME-----FNHIVPHGDM 443
QY 554 YTKVNEEH-----LCVLDRNKPLFTMPRLDRVKNLTGLVWGCKPKURE 601
Db 444 GETEETEEHPTSPDPPIWAEIMRFPSKPKPMILALARPDPCKNTITLVKAGECRPIRE 503
QY 602 LANLVVGGDRRK-BSKOLEEAKMKMPFLDKYNLNGQFRMISOMNRINV-ELRY 659
Db 504 LANLTLINGNRDGDIMSTSSVLSVTKLIDYDLVGQVAY--PKHKQADVPEIRL 561
QY 660 ICDTGAFAVQALYEAFLTVVEAMTCGLPTFATCGGPAEIIYHGSGFNIDPYHGQA 719
Db 562 AKTGVFINPAFIEPFGTLIEAAHGLPMVATNGGVPDIQVLDNGLVDPHQQSI 621
QY 720 ADILVDFEKKCKDPSHMDXISQGLKRIEKEYTMKISERLLT 763
Db 622 ATALL-----KLVDKQLMTKCOQNGLKNI-HLYSWPESHKTYLS 660

RESULT 14

US-09-394-272-7
Sequence 7, Application US/09394272
Patent No. 6472588
GENERAL INFORMATION:
APPLICANT: Haigler, Candace H.
APPLICANT: Holaday, A. Scott
TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED
FILE REFERENCE: 201304/1000
CURRENT APPLICATION NUMBER: US/09/394,272
CURRENT FILING DATE: 1999-09-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 1045
TYPE: PRT
ORGANISM: Beta vulgaris
US-09-394-272-7

Query Match 9.9%; Score 421.5; DB 4; Length 1045;
Best Local Similarity 28.1%; Pred. No. 1.1e-31;
Matches 164; Conservative 93; Mismatches 210; Indels 117; Gaps 28;
QY 241 GPTARVLEMLQLDLLEATDPCTLEKLGRIPMVFNVIITPHGYAQDNV-LGY-PD 298
Db 133 GOSTPRPLPRIN-SIDAMETWISOKEKCL-----YVLISLHGLIRGENMEIGRSD 184
QY 239 TGGGVVYILDQVRALENEMLLRIKQGLNITPRILITRLL--PVAVGTCGRLEKY-- 354
Db 185 TGGQVRYVVELARALGS-----MPGV--RVDLITRQVSSPD-VDMVSGEPTMLNP 233
QY 355 ---YGTESD-----ILRVPRTEKGIIVRKWISFEKVMPLYLETYEDVAHEI-- 399
Db 234 RDSNGFDDDDDEMSSSGAYIVRIPE---GPRDKYIAK-ELWMIYIPEFVUGALNHIYQ 288
QY 400 -SK---ELHGT-----PDIIIGNKSDGNIVASLAKHGVTOCTTAHALEKTKYPD--- 446
Db 289 MSKVLGEQIGSGGETWPAVIGHYVADAGDSALLSGGLNVPMLTGHSGRDKLEQLKQ 348
QY 447 ---SDIYWKLEDKYHSCQFTADLFAMNHTDPIITSTPQELNGSKDTGVQYESHAF 501
Db 349 GRMSKDDI---NNYKIMRIEAEUSLDASEIVITSTROEIE---EQMHLVDF 397
QY 502 TLPGH-----YRVHAGIDV---FDPKFNVISPGADMEIYFYTEERKRLKHFPRIIDL 553
Db 398 D-PVLERKLRAMKKGVCSCYGRFMRPMVYIPRGME-----FNHIVPHGDM 443
QY 554 YTKVNEEH-----LCVLDRNKPLFTMPRLDRVKNLTGLVWGCKPKURE 601
Db 444 GETEETEEHPTSPDPPIWAEIMRFPSKPKPMILALARPDPCKNTITLVKAGECRPIRE 503
QY 602 LANLVVGGDRRK-BSKOLEEAKMKMPFLDKYNLNGQFRMISOMNRINV-ELRY 659

Db 504 LANLTLINGNRDGDIMSTSSVLSVTKLIDYDLVGQVAY--PKHKQADVPEIRL 561
QY 660 ICDTGAFAVQALYEAFLTVVEAMTCGLPTFATCGGPAEIIYHGSGFNIDPYHGQA 719
Db 562 AKTGVFINPAFIEPFGTLIEAAHGLPMVATNGGVPDIQVLDNGLVDPHQQSI 621
QY 720 ADILVDFEKKCKDPSHMDXISQGLKRIEKEYTMKISERLLT 763
Db 622 ATALL-----KLVDKQLMTKCOQNGLKNI-HLYSWPESHKTYLS 660

RESULT 15

US-08-356-354-4
Sequence 4, Application US/08356354
Patent No. 5767365
GENERAL INFORMATION:
APPLICANT: SONNEMALD, Uwe
TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE
PREPARATION OF PLANTS WITH CHANGED SUCROSE CONCENTRATION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,354
FILING DATE: 20-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/EP93/01605
FILING DATE: 22-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P42 20 758.4
FILING DATE: 24-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Weillman, Edward A.
REGISTRATION NUMBER: 24,735
REFERENCE/DOCKET NUMBER: P/951-105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 382-0700
TELEFAX: (212) 382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1054 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-354-4

Query Match 9.9%; Score 419; DB 1; Length 1054;
Best Local Similarity 27.4%; Pred. No. 2e-31;
Matches 155; Conservative 93; Mismatches 207; Indels 110; Gaps 24;
QY 267 EKFLIRIMVENV-----VILTPHYGAQDNV-LGY-PDTGGQVVYIL 307
Db 139 ESTRGRLRIRISSVETMEAWGQGRGKLYIVLISLHGLIRGENMELGRDSDTGGQVYV 198
QY 308 DQVRALEN-----EMLRKQGLNITPRI-----LIIRLDPVAVGTCGQGLE 352
Db 199 ELARLGMSPGVRYVLLTRQVS-----SPVDWSYGEPTMLTPISTDGMTWNGE--- 250
QY 353 KYVGTESHDILRVPRTEKGIIVRKWISFEKVMPLYLETYEDVAHEI---SK---ELHGT 406

```
Db      251 ----SSGAYIIRIPF---GPREKYPK-EQLMPYIPFVDGALNHIIQMSKVLGEQIGS 301
Oy      407 -----PDLIIGNXSDGNIVASLAKLGVTOCTTAHALEKTKYPPSDIYMKKLEDK---- 457
Db      302 GYPWPVVAIGHYADAGSAAISGALNVPMFTGHSIGROKLEQLAQGRKSKDEINST 361
Oy      458 YHFGQFTADLFAMNHTDPIITSTFQEIAGSKDVGQYESHTAFTLPGLYRVVHG----- 512
Db      362 YKIMRRIABEELTDASEIVITSTROEID-----EQWRLYDGFD-PILERKLRARIKN 414
Oy      513 ---IDVPPKRNIVSPGADMEIYFPY-----TEKRRLKHFHPEIEDLYTKVNEEH 562
Db      415 VSCYGRFMPMAVVIIPPGEFHHIVPHEGDMGETESSEDGKTPDPPIW-----AEI 465
Oy      563 LCVLNDRNKPIIFTWPRLDVYNLTGLVEMCGKNPKLRELANLIVVGGDRRK-ESKDLEE 621
Db      466 MRFFSNPRKPMILALARPDPKKNLTLYKARGECPRLRELANTLIMGNRDNIDEMSTN 525
Oy      622 KAEMKMFELIDKYNLNGQFRWISSQMRIRNV-ELYRYICDTKGAFOVOPALYEAFGTLV 680
Db      526 SALLSLILKIMIDKYDLYGVAV--PKHHQSDVPDIYRLAAKTGVFINPAFIEPFGTL 583
Oy      681 VEMATCGIPTATCNGPAEIIYHKGSGFNIDPHYGDAQADILVDFEKKCKDPGHMDKI 740
Db      584 IEAAAYGILPMVATKXGGVDIHRVLNDGILVDPHQAIADALL---KLVAHQQLMAKC 639
Oy      741 SQGLKRIEBKTYW---KIYSERL 761
Db      640 RANGIAKNI-HLPSWPEHCKTYLSRI 663
```

Search completed: June 2, 2004, 14:53:49
Job time : 25 secs

CC genome, the seeds of the plant and fibre isolated from the plant. The
CC method is useful for altering fibre development or properties of a fibre
CC producing plant like cotton plant. Therefore, the method is useful for
CC enhancing fibre yield, enhancing fibre quality and for increasing seed
CC size in a fibre producing plant. The present sequence represents the DNA
CC encoding the cotton sucrose synthase (Susy) protein used to create the
CC transgenic plant of the invention

XX Sequence 2625 BP; 730 A; 560 C; 605 G; 729 T; 0 U; 1 Other;

Query Match 100.0%; Score 2624; DB 6; Length 2625;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGCTGAGCGTCTCTCACTCGGTCACAGTCTCCGTGAGCGTTGGATGAGACCTT 60
DB 1 ATGCTGAGCGTCTCTCACTCGGTCACAGTCTCCGTGAGCGTTGGATGAGACCTT 60
QY 61 CTTCCTCAGAGAACGAGATTTTGGCTTCTCAAGATCGAGGCAAGAAAAGA 120
DB 61 CTTCCTCAGAGAACGAGATTTTGGCTTCTCAAGATCGAGGCAAGAAAAGA 120
QY 121 ATTCTGCAACCCANTCAAATTAATCTAGAGTTGAAGCTATCCCTGAAGAGAGAAG 180
DB 121 ATTCTGCAACCCANTCAAATTAATCTAGAGTTGAAGCTATCCCTGAAGAGAGAAG 180
QY 181 AAGCTCGCTAAGTGTGCAATTTTGAAGTATGAAGCTAGTCAAGAACGATCGTGTG 240
DB 181 AAGCTCGCTAAGTGTGCAATTTTGAAGTATGAAGCTAGTCAAGAACGATCGTGTG 240
QY 241 CCTCCATGAGTGTGCACTTCTGTGTCGTCGAAGCGCTGTGTTGGAGTACATTAGTGTG 300
DB 241 CCTCCATGAGTGTGCACTTCTGTGTCGTCGAAGCGCTGTGTTGGAGTACATTAGTGTG 300
QY 301 AATGTTACCGCGCTGTGTTGAGAGAACTCACTGTTGTGAGTATCCCACTTGAAGAA 360
DB 301 AATGTTACCGCGCTGTGTTGAGAGAACTCACTGTTGTGAGTATCCCACTTGAAGAA 360
QY 361 GAGCTTGTGAGTGAAGTGAAGTGAAGAACTTGTGTTGAAATGGAATTTGAGCCCTTC 420
DB 361 GAGCTTGTGAGTGAAGTGAAGTGAAGAACTTGTGTTGAAATGGAATTTGAGCCCTTC 420
QY 421 AACTCATCATTTCCCGCCCAACTCTTCAAAATCATTGGTAATGGTGTGAGATTCCTA 480
DB 421 AACTCATCATTTCCCGCCCAACTCTTCAAAATCATTGGTAATGGTGTGAGATTCCTA 480
QY 481 AATGTCACCTTTGGGCAAAATGTTCCATGACAAAGAGACATGACCCCTTGTCTGAA 540
DB 481 AATGTCACCTTTGGGCAAAATGTTCCATGACAAAGAGACATGACCCCTTGTCTGAA 540
QY 541 TTCTCTAGAGTCAATTTGTCAAGAGGCAAGAACTGATGTTGAATGACAGAAATCAGAAC 600
DB 541 TTCTCTAGAGTCAATTTGTCAAGAGGCAAGAACTGATGTTGAATGACAGAAATCAGAAC 600
QY 601 TTGAATGCTCTTCAACATGTTTGAAGAGCAAGAGATCTTGGTACCTTACTCTCT 660
DB 601 TTGAATGCTCTTCAACATGTTTGAAGAGCAAGAGATCTTGGTACCTTACTCTCTCT 660
QY 661 GAGACACCATGTGCGAATTGGAACACCGGTTCAGAGAAATCGGTTTGGAAAGAGTGTG 720
DB 661 GAGACACCATGTGCGAATTGGAACACCGGTTCAGAGAAATCGGTTTGGAAAGAGTGTG 720
QY 721 GGTGACACCGGAGACGGGTCTCGAGATGATCAACTCTTGGATCTTCTTGAAGCA 780
DB 721 GGTGACACCGGAGACGGGTCTCGAGATGATCAACTCTTGGATCTTCTTGAAGCA 780
QY 781 ACTGATCTCTTGAACCTTGAAGAGTCTTGGAGAAATCCCAAGTGTGTTCAATGTTGTG 840
DB 781 ACTGATCTCTTGAACCTTGAAGAGTCTTGGAGAAATCCCAAGTGTGTTCAATGTTGTG 840
QY 841 ATTCTCACTCCCAAGGATTAATTCGCTCAAGACAAATGTTTGGGGTATCCCGAACCGGT 900
DB 841 ATTCTCACTCCCAAGGATTAATTCGCTCAAGACAAATGTTTGGGGTATCCCGAACCGGT 900
```

```
QY 901 GGGCAGGTTGTTACATCTTGTGATCAAGTCGAGCTTTGGAGAAATGATGCTCTCCGT 960
DB 901 GGGCAGGTTGTTACATCTTGTGATCAAGTCGAGCTTTGGAGAAATGATGCTCTCCGT 960
QY 961 ATTAAGCAAAAGACATCAACCCCTCGAATCCTCTATTATTAATCACTTCTCTCT 1020
DB 961 ATTAAGCAAAAGACATCAACCCCTCGAATCCTCTATTATTAATCACTTCTCTCT 1020
QY 1021 GATGCTGTGCGAACCAACATGCGGTCAACGACTTGAAGAAATATACGAGACACTCG 1080
DB 1021 GATGCTGTGCGAACCAACATGCGGTCAACGACTTGAAGAAATATACGAGACACTCG 1080
QY 1081 GATATTTCTGAGTACCTTCAAGAACGAAAGGAAATGTTGCAAAATGATCTCAAGA 1140
DB 1081 GATATTTCTGAGTACCTTCAAGAACGAAAGGAAATGTTGCAAAATGATCTCAAGA 1140
QY 1141 TTTGAAAAAAGTCTGCGCATCTTGGAAAACCTTACACAGAGATGTGTCTATGAAATCTCC 1200
DB 1141 TTTGAAAAAAGTCTGCGCATCTTGGAAAACCTTACACAGAGATGTGTCTATGAAATCTCC 1200
QY 1201 AAAGAGTTGACGCGACGCGAGATCTGATCATCGGAAACNACAGCGCAATATCTGT 1260
DB 1201 AAAGAGTTGACGCGACGCGAGATCTGATCATCGGAAACNACAGCGCAATATCTGT 1260
QY 1261 GCTCTCTTGTGCGACATAAATTAAGTGTCAACAAGTCAACATGCGCATGCTTTGGAG 1320
DB 1261 GCTCTCTTGTGCGACATAAATTAAGTGTCAACAAGTCAACATGCGCATGCTTTGGAG 1320
QY 1321 AAGACAAAATATCAAGTTCATATCTATTTGGAAGAAAGTTTGAAGCAATTCATTC 1380
DB 1321 AAGACAAAATATCAAGTTCATATCTATTTGGAAGAAAGTTTGAAGCAATTCATTC 1380
QY 1381 TCTTGCCAATTTACAGTGTATCTTTTTCGATGAAACCATACAGATTCATCATCACAGT 1440
DB 1381 TCTTGCCAATTTACAGTGTATCTTTTTCGATGAAACCATACAGATTCATCATCACAGT 1440
QY 1441 ACTTTCAGAGAAATTTGAGAGAACCAAGACATGTTGTCAATACGAGACCACTGCT 1500
DB 1441 ACTTTCAGAGAAATTTGAGAGAACCAAGACATGTTGTCAATACGAGACCACTGCT 1500
QY 1501 TTCACTTCTTCCGTCTCTACCGGTGTTTACATGATGATGATGATGATGATGATGATG 1560
DB 1501 TTCACTTCTTCCGTCTCTACCGGTGTTTACATGATGATGATGATGATGATGATGATG 1560
QY 1561 AACATGTTTCCCTGCGTCTGATGAGATATATCTTCCCTTACACCGAAGAGAACGG 1620
DB 1561 AACATGTTTCCCTGCGTCTGATGAGATATATCTTCCCTTACACCGAAGAGAACGG 1620
QY 1621 AAGTTGAAGATTTCCATCTGAGATGGAAGACCTTCTTACACCAAGTTGAGATGAA 1680
DB 1621 AAGTTGAAGATTTCCATCTGAGATGGAAGACCTTCTTACACCAAGTTGAGATGAA 1680
QY 1681 GAACACTTATGTGTCTCAATGACCGGACCAACCAATCTTCTTCAACAAAGGCTT 1740
DB 1681 GAACACTTATGTGTGTCTCAATGACCGGACCAACCAATCTTCTTCAACAAAGGCTT 1740
QY 1741 GATCGTGTCAAGAACTTAAACCGGACTGTGAGTGTGCGGCAAGAACCCAAAGTGTGCT 1800
DB 1741 GATCGTGTCAAGAACTTAAACCGGACTGTGAGTGTGCGGCAAGAACCCAAAGTGTGCT 1800
QY 1801 GAGTTGGCTAACCTCGTATGTTGATGATGATGATGATGATGATGATGATGATGATGAA 1860
DB 1801 GAGTTGGCTAACCTCGTATGTTGATGATGATGATGATGATGATGATGATGATGATGAA 1860
QY 1861 GAGAAAGCTGAATTAAGAAAGAAATGTTTGAAGTATGACATGACAGTACAACTTGAACGCGCAA 1920
DB 1861 GAGAAAGCTGAATTAAGAAAGAAATGTTTGAAGTATGACATGACAGTACAACTTGAACGCGCAA 1920
QY 1921 TTCAAGATGATATCATCTCAATGAAGACAGATCCGAATGTTGAATCTTACCGATACATT 1980
DB 1921 TTCAAGATGATATCATCTCAATGAAGACAGATCCGAATGTTGAATCTTACCGATACATT 1980
```

QY 1981 TGGCAGCAGAAAGTGCTTTGTATGACGCTGATGTATGAAGCTTTGATGACCTT 2040
 DB 1981 TGGCAGCAGAAAGTGCTTTGTATGACGCTGATGTATGAAGCTTTGATGACCTT 2040
 QY 2041 GTGAGGCGAATGATCTTGGGTTTGGCAACATTTCCGACCTTAAACGGTGAACCGCCAG 2100
 DB 2041 GTGAGGCGAATGATCTTGGGTTTGGCAACATTTCCGACCTTAAACGGTGAACCGCCAG 2100
 QY 2101 ATTAATGTCATGGGAAATCTGTTTCAACATGATCTTCAACATGATCAAGCTGCT 2160
 DB 2101 ATTAATGTCATGGGAAATCTGTTTCAACATGATCTTCAACATGATCAAGCTGCT 2160
 QY 2161 GACATACCTGTCGATTTCTTTGAAAGTGAAGAAAGATCCATCTCACTGGGATTAAGATC 2220
 DB 2161 GACATACCTGTCGATTTCTTTGAAAGTGAAGAAAGATCCATCTCACTGGGATTAAGATC 2220
 QY 2221 TCCCAAGAGGCTTGAAACGAATAGAGAGAAATATACATGAGATTTACTCGAGAGA 2280
 DB 2221 TCCCAAGAGGCTTGAAACGAATAGAGAGAAATATACATGAGATTTACTCGAGAGA 2280
 QY 2281 CTATTGACCTCGACAGAGATGATGATTTCTGSAAGCAATGTTTCCAACTGGAAGCCGT 2340
 DB 2281 CTATTGACCTCGACAGAGATGATGATTTCTGSAAGCAATGTTTCCAACTGGAAGCCGT 2340
 QY 2341 GAGAGTCGCTGATCTTGAAGATGTTTATGCTCTTAAGTACCTGTAAGCTGGCTGAATCA 2400
 DB 2341 GAGAGTCGCTGATCTTGAAGATGTTTATGCTCTTAAGTACCTGTAAGCTGGCTGAATCA 2400
 QY 2401 GTTCATTTGCAAGAGAGATTAATGAACCTGTTAAATACATGAGCGGTTTCTTGG 2460
 DB 2401 GTTCATTTGCAAGAGAGATTAATGAACCTGTTAAATACATGAGCGGTTTCTTGG 2460
 QY 2461 AGAATATATCTGTTTGTATGATTCATATGAGAAAGCTCTTGTATTTATCTGTCT 2520
 DB 2461 AGAATATATCTGTTTGTATGATTCATATGAGAAAGCTCTTGTATTTATCTGTCT 2520
 QY 2521 TTTCCTTTCTTTTTCGCGGCAATGTTGAACATGGGTTGCGCCGCTCAATCC 2580
 DB 2521 TTTCCTTTCTTTTTCGCGGCAATGTTGAACATGGGTTGCGCCGCTCAATCC 2580
 QY 2581 AGTTAATATGATGATCTTTGTTTCAAAAAA 2625
 DB 2581 AGTTAATATGATGATCTTTGTTTCAAAAAA 2625

RESULT 2
 ID AAA67143 standard; DNA; 2906 BP.
 AC AAA67143;
 XX 31-OCT-2000 (first entry)
 DE Eucalyptus grandis sucrose synthase nucleotide sequence SEQ ID NO:139.
 KW Eucalyptus grandis; pinus radiata; Monterey pine; modification;
 KM plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;
 XX transgenic plant; ds.
 OS Eucalyptus grandis.
 XX MOJ00022092-A2.
 PN 20-APR-2000.
 PD 08-OCT-1999; 99MO-N2000169.
 PF 13-OCT-1998; 98US-00170862.
 PR 11-AUG-1999; 99US-0148426P.
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX

PI Blocksberg LN;
 XX MPI: 2000-339328/29.
 DR P-PSDB; AAB16336.
 DR New genes encoding proteins involved in a plant polysaccharide
 PT biosynthetic pathway, useful for modulating or altering the
 PT polysaccharide content, composition or structure of the plant.
 XX Claim 1; Page 104-105; 301pp; English.
 XX The present invention describes isolated polynucleotides (PN) comprising
 CC a sequence selected from one of 835 nucleotide sequences given in
 CC AAA67073 to AAA67907, their (reverse) complements, sequences producing an
 CC Expectation (E) value of 0.01 or less compared to the 835 sequences,
 CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
 CC 835 sequences or sequences that are degenerately equivalent or allelic to
 CC the 835 sequences. The polynucleotides are used to modify the activity of
 CC a polypeptide involved in a polysaccharide biosynthetic pathway in the
 CC plant. They are especially used to modulate or alter the polysaccharide
 CC content, composition or structure of the plant. AAB16268 to AAB16340 are
 CC proteins encoded by some of the polynucleotide sequence given in the
 CC present invention
 CC
 SQ Sequence 2906 BP; 679 A; 709 C; 731 G; 787 T; 0 U; 0 Other;
 XX
 Query Match 57.1%; Score 1499.4; DB 3; Length 2906;
 Best Local Similarity 76.5%; Pred. No. 0;
 Matches 1852; Conservative 0; Mismatches 567; Indels 3; Gaps 1;
 QY 1 ATGGCTGAGCGTGTCTCATCTGCGTCACAGTCTCCGTGACGTTGATGAGACCTT 60
 DB 233 ATGGCTGAGCGTGTCTCATCTGCGTCACAGTCTCCGTGACGTTGATGAGACCTT 60
 QY 61 CTCTCTCAGAGAAAGATTTG3CTTCTCTCAAGATGAGGCAAGAAAGAA 120
 DB 293 TCTCTCTCAGAGAAAGATTTG3CTTCTCTCAAGATGAGGCAAGAAAGAAAG 352
 QY 121 ATTCGCAACCATCAATATCTAGATTTAAGTATCCCTGAAGAAAGAAAG 180
 DB 353 ATCTTGACGCGCCACAGATTTTGTAGATTTAAGTATCCCTGAAGAAAGAAAG 412
 QY 181 AAGCTCCGTAATGATGATTTTGAAGTATGAAGCTATGACGAAGCATGCTT 240
 DB 413 AAGCTCCGTAATGATGATTTTGAAGTATGAAGCTATGACGAAGCATGCTT 240
 QY 413 AAGCTCCGTAATGATGATTTTGAAGTATGAAGCTATGACGAAGCATGCTT 240
 DB 241 CTCTCATGAGTGTGACCTTGTCTGTCGCAAGGCTG3GTTTGGAGTACATTAAGTG 300
 DB 473 CTCTCATGAGTGTGACCTTGTCTGTCGTCGCAAGGCTG3GTTTGGAGTACATTAAGTG 300
 QY 301 AATGTTACGCGCTTGTGTTGAGAACTCATGTTGCTGATATCTTCAAGAA 360
 DB 533 AAGCTCAATGCGCTGTTGTTGAGAACTCATGTTGCTGATATCTTCAAGAA 592
 QY 361 GAGCTTGATGAGAACTTCAATGAAGAACTTTGTTGAAATGAAATTTGAGACCTT 420
 DB 593 GAGCTTGATGAGAACTTCAATGAAGAACTTTGTTGAAATGAAATTTGAGACCTT 420
 QY 421 AACTCATATTTCCCGCCCAACTTTTCAAAATCAATGGAATGTTGAGTCTTA 480
 DB 653 AACTCATATTTCCCGCCCAACTTTTCAAAATCAATGGAATGTTGAGTCTTA 480
 QY 481 AATGTCATCTTCCGCTTAAGCTTTTCATGACAAAGAAAGCTTCACTTCTGAA 540
 DB 713 AATGTCATCTTCCGCTTAAGCTTTTCATGACAAAGAAAGCTTCACTTCTGAA 540
 QY 541 TTCTCAGAGTGTGTCACAAAGGCAAGAACTATGTTGAATGACAGATTCAGAAC 600
 DB 773 TTCTCAGAGTGTGTCACAAAGGCAAGAACTATGTTGAATGACAGATTCAGAAC 600
 QY 601 TTGAATGCTTTCAACATGTTTGAAGAAAGCAAGAGATGTTGATCTTCTCT 660
 DB 833 GTGTTCTCCCTCAACATGTTTGAAGAAAGCAAGAGATGTTGATCTTCTCT 660

QY	661	GAGACACCACTGATGCCGAATTCGAA	CACCGGTTCCAGAGAAATCGGTTGGAAAGAGGTTGG	720
Db	893	GAGACCCCGTACTCCAGTTCTGAGCA	CAAGTTCCAGAGAGATCGGACTCGAGCGGGGGTGG	952
QY	721	GGTACACCCGAGAACGGCTGCTGAGAT	GTGATCCAACTCCCTTTTGGATCTCTTGAGGACA	780
Db	953	GGTACACCGCTGAGCGGCTCTGAGAT	GTATCCAGCTCTCTGTTGGATCTCTTGABGGCT	1012
QY	781	ACTGATCCCTTGGACCTTTGAGAA	GTTCTCTTGGAGAAATCCCAATGCTGTCAATGTTGTG	840
Db	1013	CCCGACCCGTCACCTCTGAGAA	GTTCTTGAGATAGGGTTCCCATGAGCTTCAAGTGTG	1072
QY	841	ATTCTCACTCCCAACGGAATCTTGCT	CAACAATGTTTGGGGTATCCCGAACCGGCT	900
Db	1073	ATCATGTCTCCCAACGGAATCTTGCT	CAAGACAGCTCTTGGTGTATCCGGAATACCGGT	1132
QY	901	GGGCAAGGTTGTTACATCTTGAT	CAAGTCGGAAGCTTTGGAGAAATGAGATGCTCTCCGT	960
Db	1133	GGCAAGGTTGTTACATCTTGATCA	AGTCAGTGTGTGTCCTTGAAGAAAGAAATGCTTCAACGC	1192
QY	961	ATTAAGCAACAGGACTCAACATCA	CCCCCTTGAAATCTCATTTATTAACATCTTCTT	1020
Db	1193	ATTAAGCAACAGGACTGATATTA	ACTCTCTGGAATCTCATTTATTAACATCCGGCTTCTTCA	1252
QY	1021	GATGCTGTCCGAAACAATGCGGTC	ACAAGCATTTGAGAAAGATATACGGAACAGACAATCG	1080
Db	1253	GACCGGTTGGAACCACTGTGGC	AGCGCTTGAGAAAGTTTTGGAGCCGAATACCTC	1312
QY	1081	GATATCTTCGAGTACCTTCAGAA	CAGAAAGGGAATTTGTCGAAATATGANTCTCAGA	1140
Db	1313	CACATCTTCGCGCTCCCTTCAG	AATGAAAGAGAGGAGTGTCCCGAGTGTTCGGG	1372
QY	1141	TTTGAAAAAGTCTGSCCAT	CTTGGAACCTTACACAGAGATGTTGCTCATGAATCTCC	1200
Db	1373	TTTCG---AGGTGTGGCCCTAT	TTTGGAAAGATACCTGAGATGTGCGAGCGGAATTTGCT	1429
QY	1201	AAAGAGTTGCACGGCAACGCCA	GAATCTGATCATGTGGAAACUNACACGACCGCAATATGTC	1260
Db	1430	GGAAGTTGCAAGGCGACGACCT	GTATCTGTATTCGGAACCTACAGTATGAGAACCAATGT	1489
QY	1261	GCTCTCTTGTCTCGACATAA	ATTTAGGTGTCAACAGTGCACCATCGCCATGCTTTGGAG	1320
Db	1490	GCTTCTCTGTAGCACATTA	ATTTAGGTGTTCACAGTGTACAAATAGCCCATGGCCTCGAG	1549
QY	1321	AAGAACAANAATCCAGATT	CAGATCTTATTGGAAAGAACTTTGAAGACAATTACATTT	1380
Db	1550	AAGACGAAGTACCGAGAGT	CAACAATAATCTGGAAGAAATTTGAGAAAGATACACATTC	1609
QY	1381	TCTTGCCCAATTTACAGCT	GAATCTTTTGGCAATGAAACCATACAGATTCATGATCACAGT	1440
Db	1610	TCTTGCCCAATTTACAGCTGA	ATCTATCTCCATGAAACACACCGACCTTCATTTATCACAGC	1669
QY	1441	ACTTTCAGAAATTTGACAG	AAGCAAGCACTGTTGTCAATACGAAGCCACACTGCT	1500
Db	1670	ACCTTCCAGAAATTTGCTG	GAAGCAAGATACGTGGGCAGTATGAGAGTCAATGAAC	1729
QY	1501	TTCACTGTTCTGCTGCTT	ACCGGTGTTCATGTGTATTCGATGCTTTGATCCCAATTC	1560
Db	1730	TTCACTGTTCTGCACTT	ACCGAGTTTCCACGAGATCGACGTCCTTCGACCCCAAGTTC	1789
QY	1561	AACAATGTTTCCCCCGT	GTGATATGAGADATATCTCCCTTACACCGAAGAGAAAGCGG	1620
Db	1790	AACAATGTTTCAACGAG	GTGCTGACATGACATCTTCTGTTACACCGAACAGAGCGG	1849
QY	1621	AGGTTGAAGCATTTCCAT	CTTGAGATCGAAGCTTCTTTTACACCAAGTTGAGAAATGA	1680
Db	1850	CGGTTGAAGTCTTCCAC	CTCGAGATCGAGAAAGTCTTCAAGGAGTGTGAAGAACAG	1909
QY	1681	GAAACATTATGTGTGCT	CAATGACCGCAACAAGCAATTCGTGTTCAAAATGCCAAGCTT	1740
Db	1910	GAAACCTGTGTGTGTT	GAAAGATTAAGAAAGAGCTATTAATTTTTCACCATGGAAGGCTG	1969

Qy	1741	GATCGTGTCAAAGAACTTAAACCGGACTCGTCGAGTCGGGCGGCAAGAACCCAAAGTTGGCT	1800
Db	1970	GACCGTGTCAAAGAACTTGAACAGGACTTGTGTGAGTGGATGATGGCAAGAACTCCAGTTGAGC	2029
Qy	1801	GAGTTGGCTAACTCCGTGATGTTGTAGGTGGTATAGCGCAAGAAATCTAAAGATTTTGAA	1860
Db	2030	GAACCTGCCCAACTTTGGTCGTGTGTGAGGTGACAGAGAGAAAGGATTCGAAGGACTTGGAA	2089
Qy	1861	GAGAAAGCTGAAAATGAAAGAAAATGTTTGTAGCTGATGCAAGTACAACTTGAACGSCCAA	1920
Db	2090	GAGCAGCTGTAGATGAAGAAAATGTACGACCTCATCGAAAATGATCAAGCTGAATGGCCAG	2149
Qy	1921	TTCCAGATGGAATATCATCTCAAATGGAACAGAAATCCGAAATGTTGAACCTTACCGATACAT	1980
Db	2150	TTCAAGGTGGAATTTCTCTCCAGATGAACCGGGTGAAGAAATGGAAGGCTCTACCGCTACATC	2209
Qy	1981	TGCGACACGAAGAGGTGCTTTGTATACAGCTCGTCACTGTATAGAGCCTTGTGATTGACAGTT	2040
Db	2210	TGTACACGAAGGAGAGTCTTTCGTTCCAAACGGCTATCTATGAGCCTTTCCGGTGTGACCGTG	2269
Qy	2041	GTGGAGGCAATGACTTGGCGGTTTGGCCAAACATTGCAACTGTAAACGGTGAACCAAGCCGAG	2100
Db	2270	GTTGAGGCCATGACTGTGTGATTTGCCAACTTTGGCCACTTGGCAATGGTGAACCAAGCTGAG	2329
Qy	2101	ATTATTTGTCCATGAGGAATCTGTTTTCACATTTGATCTTTACCAATGGTATCAAGCTGCT	2160
Db	2330	ATCATTTGTGATGACCAAAATCGGGCTACCACTTGATCTCTATGACATGGTGAACAGGCGCC	2389
Qy	2161	GACATCTGTCGCAATTTCTTTGAAAAGTGTAAAGAAATCATCTCACTGGGAGTAAAGTC	2220
Db	2390	GAGCTTCTTTGAGCTTCTTCACACATGTCAGAGTTGACAGATGCCACTGGAGCAGAGATC	2449
Qy	2221	TCCCAAGAGAGCTTTGAAACGAAATAGAGAGAGATATCATGAGAAATTTACTCGGAGAGA	2280
Db	2450	TCAAGGGTGCCATGACAGAGAAATTGAAGAGAAATATCATGAGAAAATATATTTCTGAAGG	2509
Qy	2281	CTATTTGACCTGTACAGAGATGTATGATTTCTGGAAGCATGTTTCCAACTTTGAACGCCGT	2340
Db	2510	CTGTTGAACCTGACTGCGCGTATGGCTTCCTGGAAGCATGTGACTAACTTGTATCGCGCGC	2569
Qy	2341	GAGAGTGTGTGTAACCTTGAGATGTTTATGCTTTATAGTACGCTAAGCTGGCTGAATCA	2400
Db	2570	GAGGTGTGCGGTATCTTGAATGTTCTATGCTCTCAAGTATGCGCCACTGACAGACTCT	2629
Qy	2401	GTTCATTGGCCAGAGAGATAA	2422
Db	2630	GTTCCTCCGGCTGTGCAGTAA	2651
RESULT 3			
AAA67087			
ID	AAA67087	standard; DNA; 2913 BP.	
XX	AAA67087;		
NC			
XX			
DT	31-OCT-2000	(first entry)	
DE			
XX	Eucalyptus grandis sucrose synthase nucleotide sequence SEQ ID NO:15.		
XX			
KW	Eucalyptus grandis; pinus radiata; Monterey pine; modification;		
KW	plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;		
KW	transgenic plant; ds.		
XX			
OS	Eucalyptus grandis.		
XX			
PN	W0200022092-A2.		
PD	20-APR-2000.		
XX			
PF	08-OCT-1999;	99NC-NZ000169.	
XX			
PR	13-OCT-1998;	98US-00170862.	
PR	11-AUG-1999;	99US-0148426P.	

XX (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Bloksberg LN;
XX
DR WPI: 2000-339328/29.
DR P-PSDB; AAB16282.
XX
PT New genes encoding proteins involved in a plant polysaccharide
PT biosynthetic pathway, useful for modulating or altering the
PT polysaccharide content, composition or structure of the plant.
XX
PS Claim 1; Page 43-44; 301pp; English.
XX
CC The present invention describes isolated polynucleotides (PN) comprising
CC a sequence selected from one of 835 nucleotide sequences given in
CC AAB67073 to AAB67907, their (reverse) complements, sequences producing an
CC expectation (E) value of 0.01 or less compared to the 835 sequences,
CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
CC 835 sequences or sequences that are degenerately equivalent or allelic to
CC the 835 sequences. The polynucleotides are used to modify the activity of
CC a polypeptide involved in a polysaccharide biosynthetic pathway in the
CC plant. They are especially used to modulate or alter the polysaccharide
CC content, composition or structure of the plant. AAB16268 to AAB16340 are
CC proteins encoded by some of the polynucleotide sequence given in the
CC present invention
XX
SQ Sequence 2913 BP; 685 A; 709 C; 732 G; 787 T; 0 U; 0 Other;
Query Match 57.1%; Score 1497.8; DB 3; Length 2913;
Best Local Similarity 76.4%; Pred. No. 0;
Matches 1851; Conservative 0; Mismatches 568; Indels 3; Gaps 1;
QY 1 ATGCTGAGCGTCTCTCACTCGCTCCAGCTCTCGTGAAGCGTTGGATGAGACCTT 60
DB 242 ATGCTGAGCGTCTCTCACTCGCTCCAGCTCTCGTGAAGCGTTGGATGAGACCTT 301
QY 61 CTCTGCTCAGAGAGAGATTTTGGCTTGTCTCAAGGATCGAGGGAAGAAAGGA 120
DB 302 TCTCTCAGCGCAAGATATTTGGCTTGTCTCAAGGATCGAGGGAAGAAAGGA 361
QY 121 ATTCTGCAACACATCAATATTATCTAGATTTGAAGCTATCCCTGAGAGAAAGAAAG 180
DB 362 ATCTTGACGCGCCACACATTTTCTGAGATTTGAGGCACTCTGAGAGAGAGAGACA 421
QY 181 AAGCTCGCTAATGTGATTTTGAAGATTTGAAGGCTAGTCAGAGAGAGATGCTGTG 240
DB 422 AAGCTCTTGAATGGGCTTTGGTGAAGTCTCAAAATCAGCTCAGAGAGAGATGCTG 481
QY 241 CCTCAGAGGTTGCACTTGTCTGTCTGTCAGAGCGTGTGTTGGAGATCATTAAGTGTG 300
DB 482 CCTCAGAGGTTGCTTGTCTGTCTGTCAGAGCGGCGTGTGGAACACATCCGTGTG 541
QY 301 AATGTTCAAGCGCTTGTGTTGAGAGAACTCACTGTGAGTATCTCCACTTCAAGGA 360
DB 542 AAGCTTCATGCGCTTGTGTTGAGAGAACTCACTGTGAGTATCTCCACTTCAAGGA 601
QY 361 GAGCTTGTGAAGAAAGTTCAATGGAACCTTGTGTTGAATGATTTTGAAGCCTTC 420
DB 602 GAGCTTGTGAAGAAAGTTCAATGGAACCTTGTGTTGAATGATTTTGAAGCCTTC 661
QY 421 AAGCTATATTTCCCGCGCGCACTTTTCAAGTCTATTTGCAATGAGGCTGATTTCTC 480
DB 662 AAGCTATATTTCCCGCGCGCACTTTTCAAGTCTATTTGCAATGAGGCTGATTTCTC 721
QY 481 AATGTCACCTTTGGGCAAAATTTGTCATGACAAGAGAGATGACCTTTGCTCGAA 540
DB 722 AATGTCACCTTTGGGCAAAATTTGTCATGACAAGAGAGATGACCTTTGCTCGAA 781
QY 541 TTCTCTCAAGTTCATTTGTCACAAGGCAAGAAATGATGTTGAATGACAGATTCAGAA 600
DB 782 TTCTCTCAAGTTCATTTGTCACAAGGCAAGAAATGATGTTGAATGACAGATTCAGAA 841

QY 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGAGAGATCTTGTATACCTTCT 660
DB 842 GTGTTCTCCCTCCAAACATGCTCTGAGAGAGAGAGATCTGACCTGCTCAAAACC 901
QY 661 GAGACACCATGTCGAAATTCGAACACCGGTTCCAGAAATCGGTTTGAAGAGCTTG 720
DB 902 GAGACACCGGTAATCCAGTTGAGACACAAGTTCCAGAGAGATCGGAGCTGAGCGGGGTG 961
QY 721 GGTACACCGGAGAAACGGTGTCTGAGATGATTCACCTTTTGGATCTTGAAGCA 780
DB 962 GGTACACCGGAGAAACGGTGTCTGAGATGATTCACCTTTTGGATCTTGAAGCA 1021
QY 781 ACTGATCTTCAACCTTTGAGAAATCTTGGAGAAATCCCATGATGTTCAATGTTGTG 840
DB 1022 CCGACCGGTCATCTCTGAGAAATTTTGGATGAGTTCCCATGATGTTCAAGTGTG 1081
QY 841 ATTCTCACTCCCAAGAGATCTTGTGCTCAGACATGTTTGGGATATCCGACACCGGT 900
DB 1082 ATCATGCTCCCAAGAGATCTTGTGCTCAGACACGTCCTTGGTATCCGATACCGGT 1141
QY 901 GGGCAGGTTGTTTACATCTTGAATCAAGTCCGAGCTTGGAGATGAGATGCTCTCGT 960
DB 1142 GGGCAGGTTGTTTACATCTTGAATCAAGTCCGAGCTTGGAGATGAGATGCTCTCGT 1201
QY 961 ATTAAGCAAGAGACTCAATCATCACCCCTGAAATCTCATTTATTAAGACTTCTTCT 1020
DB 1202 ATTAAGCAAGAGACTCAATCATCACCCCTGAAATCTCATTTATTAAGACTTCTTCT 1261
QY 1021 GATGCTGTGAGAACATGCGGTCAACGACTTGAAGAAATGATACGAAACAGAGACTCG 1080
DB 1262 GATGCTGTGAGAACATGCGGTCAACGACTTGAAGAAATGATACGAAACAGAGACTCG 1321
QY 1081 GATATTTCTGAGATCTTCAAGACAGAAAGGAAATGTTTCCAAATGATCTCAAGA 1140
DB 1322 CACATTTCTGCGCTCCCTTCAAGAAATGAGAGAGAGTCCGCAATGATTTCCGG 1381
QY 1141 TTTGAAAAGCTGAGCAATCTTGAAGAACTTACAGAGATGTTGCTCATGAATCTCC 1200
DB 1382 TTCC--AGTGTGAGCTTATTTGAAAGATCACTGAGATGTCGAGAGAACTTCT 1438
QY 1201 AAGAGTTGACAGGACGCGCATCTGATCATCGGAAACAAACAGCGAGCAATATCTC 1260
DB 1439 GAGAGTTGACAGGACGCGCATCTGATCATCGGAAACAAACAGCGAGCAATATCTC 1498
QY 1261 GCTTCCTTGTGCAATTAATAGTGTCAACAGTACACATGCTTGTGAG 1320
DB 1499 GCTTCCTTGTGCAATTAATAGTGTCAACAGTGTCAACATGCTTGTGAG 1558
QY 1321 AAGCAAAATATCAGATTCATATCTATTTGAAGAACTTGAAGCAAAATACATTTT 1380
DB 1559 AAGCAAAATATCAGATTCATATCTATTTGAAGAACTTGAAGCAAAATACATTTT 1618
QY 1381 TCTTCCAAATTAAGCTATCTTTTGAAGAACTTGAAGCAAAATACATTTT 1440
DB 1619 TCTTCCAAATTAAGCTATCTTTTGAAGAACTTGAAGCAAAATACATTTT 1678
QY 1441 ACTTTCAGAAATTTGAGAGAGAGACACTGTTGTCAATACGAGAGCAACTGCT 1500
DB 1679 ACTTTCAGAAATTTGAGAGAGAGACACTGTTGTCAATACGAGAGCAACTGCT 1738
QY 1501 TTCACTGCTGCTCTACCGTGTGATCATGATGATGATGATGATGATGATGATGATG 1560
DB 1739 TTCACTGCTGCTCTACCGTGTGATCATGATGATGATGATGATGATGATGATGATG 1798
QY 1561 AACATGTTTCCCTGCTGATATGAGATATATCTTCAACCGAAGAGAGCGG 1620
DB 1799 AACATGTTTCCCTGCTGATATGAGATATATCTTCAACCGAAGAGAGCGG 1858
QY 1621 AGGTTGAAGATTTTCATCTGAGATGAGAGACTTTTACACCAAGTTGAAGTGA 1680
DB 1859 CGGTTGAAGATTTTCATCTGAGATGAGAGACTTTTACACCAAGTTGAAGTGA 1918


```

Db 1107 GAGACCCCGTACTCCAGATCGAGCAAGTTCCAGAGATCGGGCTGAGCGGGGGTGG 1166
Qy 721 GGTACACCCGAGAACGGGTGCTGAGATGATCCAACTCTTTGGATCTTTGAGGCA 780
Db 1167 GGTACACGGCTGAGCGGGCTCTGAGATGATCCAGATCTCTTGGATCTCTTGAAGCT 1226
Qy 781 ACTGATCTTGCACCTCTGAGAGAGTTCTTGGAGAGATCCCATGAGTTCATATTTG 840
Db 1227 CCGACCCGTGACCTCTGAGAGAGTTCTTGGATGAGGTTCCCATGAGTTCATATCTG 1286
Qy 841 ATCTCACTCCCAAGAGTACTTGTCTGAGAGCAATGTTTTGGGTATCCGACACCGGT 900
Db 1287 ATCATGTCTCCCAAGAGATCTTGTCTGAGAGCACTCTTGGTTATCCGATACCGGT 1346
Qy 901 GGGCAGGTTGTTACATTTGGATGATCCGAGTTGGAGATGAGATGCTCTCCGT 960
Db 1347 GGGCAGGTTGTTACATCTGATGATGATGCTGAGAGAGTTGGAGAGAGATCTTCAACGC 1406
Qy 961 ATTAAGCAACAGAGCTCAACATCACTCCGATCTCATTAATTAATGATCTTCTCT 1020
Db 1407 ATTAAGCAACAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1466
Qy 1021 GATGCTGTCGAGCAACATGCGGTCAACGACTTGAAGAGATACGAGACAGAGCACTG 1080
Db 1467 GACGCGGTTGAGAACACCTGTGGCAGCGCTTGAAGAGTTTGGAGCCGAGATCTCC 1526
Qy 1081 GATATTTCTGAGATCTCTTGAAGAGAGAGATGTTTGGAGATGATCTTCAAGA 1140
Db 1527 CACATTTCTGCGCTCTCTTGAAGATGAGAGAGATGATGATGATGATGATGATGATGAT 1586
Qy 1141 TTGAGAGAGCTGAGCACTTGAAGAGAGATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1587 TTGC--AGGTGTGCTTATTTGAGAGATGATGATGATGATGATGATGATGATGATGAT 1643
Qy 1201 AAGAGTTGACCGGACCGCAGATCTGATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAT 1260
Db 1644 GAGAGTTGACGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1703
Qy 1261 GCTTCTTGTGTCGACATTAATGAGTGTCAACAGTGCACCATGCGCCATGCTTGGAG 1320
Db 1704 GCTTCTTGTGTCGACATTAATGAGTGTCAACAGTGTCAACATGAGCCATGCTTGGAG 1763
Qy 1321 AAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Db 1764 AAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1823
Qy 1381 TCTTGCCAAATTAAGAGATGATCTTTTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAT 1440
Db 1824 TCTTGCCAAATTAAGAGATGATCTTTTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAT 1883
Qy 1441 ACTTTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1884 ACTTTCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1943
Qy 1501 TTTACTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1944 TTTACTCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2003
Qy 1561 AACATGTTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Db 2004 AACATGTTTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2063
Qy 1621 AGGTGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 2064 CGGTGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2123
Qy 1681 GAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 2124 GAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2183
Qy 1741 GATGCTGTCAGAGATTAACGAGATCTGATGATGATGATGATGATGATGATGATGATGAT 1800

```

```

Db 2184 GACCGTGTCAAGAGATTTGACAGGCTTGTGATGATGATGATGATGATGATGATGATGAT 2243
Qy 1801 GAGTTGCTAACCTCGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 2244 GAACTCGCAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2303
Qy 1861 GAGAGGCTGAGATGAGAGAGATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1920
Db 2304 GAGAGGCTGAGATGAGAGAGATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 2363
Qy 1921 TTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
Db 2364 TTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2423
Qy 1981 TTGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
Db 2424 TTGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2483
Qy 2041 GTGAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Db 2484 GTGAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2543
Qy 2101 ATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db 2544 ATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2603
Qy 2161 GACATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Db 2604 GACATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2663
Qy 2221 TCCAGAGAGGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
Db 2664 TCCAGAGAGGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2723
Qy 2241 CTATGAGAGGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2300
Db 2724 CTATGAGAGGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2783
Qy 2341 GAGAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
Db 2784 GAGAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2843
Qy 2401 GTTCCATTTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
Db 2844 GTTCCATTTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2903

```

RESULT 5
ABK17072
ID ABK17072 standard; cDNA; 3103 BP.

ABK17072;
26-MAR-2002 (first entry)

Buclalypus grandis promoter polynucleotide #28.

XX Promoter; pine, leaf, root; flower; pollen; bud; meristem; xylogenesis;
XX temporally regulated promoter; Pinus radiata; Buclalypus grandis; ss;
XX PCR primer.

XX Os
XX Eucalyptus grandis.

XX W0200198485-A1.

XX 27-DEC-2001.

XX 20-JUN-2001; 2001MO-N2000115.

XX 20-JUN-2000; 2000US-00598401.

XX 28-NOV-2000; 2000US-00724624.

XX (GENE-) GENESIS RES & DEV CORP LTD.

PA (FLET-) FLETCHER CHALLENGE FORESTS IND LTD.
XX
XX Perera R, Rice S, Eagleton C, Lasham A;
XX MPI: 2002-114583/15.
DR P-PSDB; AAU80759.
XX
XX Novel polynucleotide promoter sequences from pine and Eucalyptus useful
PT for modifying expression of endogenous and/or heterologous
PT polynucleotides in transgenic plants.
XX
XX Claim 1; Page 76-77; 121pp; English.
XX
XX The invention relates to isolated promoter sequences from *Pinus radiata*
CC and *Eucalyptus grandis*, comprising a sequence chosen from leaf-, root-,
CC flower-, pollen-, bud-, meristem-specific promoters or temporally
CC flower-, pollen-, bud-, meristem-specific promoters. The promoter
CC regulated promoters such as xylome-specific promoters. The promoter
CC polypeptides and their related polynucleotides are useful in the
CC production of genetic constructs, used for modifying gene expression in a
CC target organism, in particular a plant. The method is useful for
CC modifying expression of a polynucleotide that comprises an intron
CC sequence, through removal of the intron sequence. The method is useful
CC for modifying growth and development of plants, and cellular responses to
CC external stimulus, such as environmental factors and disease pathogens.
CC The sequences are useful in genome and physical mapping, in positional
CC cloning of genes, in various assays to determine biological activity, to
CC raise antibodies, to isolate corresponding interacting proteins and other
CC compounds, and to quantitatively determine levels of interacting proteins
CC or other compounds. Sequences ABK17016-ABK17125 represent *Pinus radiata*
CC and *Eucalyptus grandis* polynucleotides and PCR primers used in the method
CC of the invention
XX
XX Sequence 3103 BP; 756 A; 740 C; 772 G; 835 T; 0 U; 0 Other;
SQ
Query Match 57.1%; Score 1497.8; DB 6; Length 3103;
Best Local Similarity 76.4%; Pred. No. 0;
Matches 1851; Conservative 0; Mismatches 568; Indels 3; Gaps 1;
QY 1 ATGGCTGAGCGTCTCTCACTCGCTCCAGCTCTCCGTGAGCGTTTGATGAGACCTT 60
DB 447 ATGGCTGATCCGATGTTGATCGAAGCCAGCGCTTGGAGAGAGACCTTC 506
QY 61 CTGTGTCAGAGAGAGATTTGGCTTGTCTCAAGAGATCGAGGCGAAGAGAGAG 120
DB 507 TCTGCTCAGCCCAAGATATGCTGCTCTTCAAGGCTTGAAGCCAGGCGAAGAG 566
QY 121 ATTGCGAAGCAATCAATATTCTAGATTTGAAGTATCCCGAAGAGAGAGAGAG 180
DB 567 ATCTTGACGCGCCACCAATATTTGCTAGATTTGAGGCCATCTTGAGAGAGAGAG 626
QY 181 AAGCTCGCTAAATGATATTTTGAAGATTTGAAGGCTAGTCAAGAGAGAGAGAG 240
DB 627 AAGCTTCTTGAATGGGGCTTTGGTGAATGCTCAAAATCAGCTCAAGAGAGAGAG 686
QY 241 CCTGATAGGTTGCACTTGTCTGTTCAGAGGCTGTGTTGGAGATCAATTAAGATG 300
DB 687 CCTCATAGGTTGCTCTTGTCTGTTCAGAGGCTGTGTTGGAGATCAATTAAGATG 746
QY 301 AATGTTACGCGCTTGTGTTGAGAGATCACTGTTGCTGATATCTCACTTCAAGGA 360
DB 747 AAGCTGCAATGCGCTTGTGTTGAGAGATCACTGTTGCTGATATCTCACTTCAAGGA 806
QY 361 GAGCTTGTGATGAGATTTCAATGAGAACTTGTGTTGAGATTTGAGATTTGAGAG 420
DB 807 GAGCTTGTGATGAGATTTCAATGAGAACTTGTGTTGAGATTTGAGATTTGAGAG 866
QY 421 AACTCATATTCCCGCCCACTTTTCAAAATCATTGTAATGATGATGATTTCTTA 480
DB 867 ACTGCTCTTTTCCGCGCCGCACTTTTCAAGTCTATTGCAATGAGGATGATTTCT 926
QY 481 AATGTCACCTTTGGGCAAAATGTTTCATGCAAGAGAGAGATGACCTTTGGTGA 540
DB 927 AATGCGCATCTCTCGCTAAGCTTTTCATGCAAGAGAGAGAGAGAGAGAGAGAG 986

QY 541 TTCTCAGAGTCCATTTGTCAAGAGGCGAAGAACATGATGTTGAATGACAGATTCAGAAC 600
DB 987 TTCTCAGAGTCCATTTGTCTCAAGAGGCGAAGAACATGATGTTGAATGACAGATTCAGAAC 1046
QY 601 TTGAATGCTCTTCAATGTTTGAAGAGAGAGAGATCTTGTACCTTACCTTCT 660
DB 1047 GTTGTCTCTTCCCAACATGTTCTGAGAGAGAGAGAGATCTGACCTTCAAAACC 1106
QY 661 GAGACACCATGTCGCGAATTTGAAACCGGTTCCAGAGAAATCGGTTTGAAGAGTTGG 720
DB 1107 GAGACCCCGTACTTCCAGTTTCAAGAGAGAGAGAGATTCAGAGAGATTCGAGCGGGGTTG 1166
QY 721 GTTGAACCGGAG 780
DB 1167 GGTGACACGCGTGAAG 1226
QY 781 ACTGATCTTGAACAG 840
DB 1227 CCGAGCCGTCAGCTTCAAG 1286
QY 841 ATTCTCACTCCCGAGATCTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 1287 ATCATGCTTCCCGAGATCTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1346
QY 901 GGGCAGGTTGTTATCATCTTGAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 1347 GGGCAGGTTGTTATCATCTTGAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1406
QY 961 ATTAAG 1020
DB 1407 ATTAAG 1466
QY 1021 GATGCTGTCGAG 1080
DB 1467 GATGCTGTCGAG 1526
QY 1081 GATGCTGTCGAG 1140
DB 1527 GATGCTGTCGAG 1586
QY 1141 TTGGAAG 1200
DB 1587 TTGGAAG 1643
QY 1201 AAG 1260
DB 1644 GAG 1703
QY 1261 GCTCTCTGTCGAG 1320
DB 1704 GCTCTCTGTCGAG 1763
QY 1321 AAG 1380
DB 1764 AAG 1823
QY 1381 TCTTGCATTTTCAAGTATCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 1824 TCTTGCATTTTCAAGTATCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1883
QY 1441 ACTTTCAG 1500
DB 1884 ACTTTCAG 1943
QY 1501 TTCACTCTTCTGCTCTCAAG 1560
DB 1944 TTCACTCTTCTGCTCTCAAG 2003
QY 1561 AACTTGTTCCTCTGCTCTCAAG 1620
DB 2004 AACTTGTTCCTCTGCTCTCAAG 2063

QY 1621 AGTTGAAGCATTTCCATCTGAGATCGAAGACCTTCTTTTACACCAAGTTGAGAAATGA 1680
 DB 2064 CGGTTGAATCTCTTCCATCTGAGATCGAAGAACTCTTCCAGGATGTTGAGAAACAAG 2123
 QY 1681 GAAACATTATGTGTCTCAATGACCGCAACAACCAATTCTGTTTCAATGCCAAAGCTT 1740
 DB 2124 GAACACTGTGTGTGTGTAAGATGAAGAAAGCCTATTATTTTCCATGCGCAAGGCTG 2183
 QY 1741 GATGTGTCAAGAACTTAAACCGGACTCGTCAGTGTGTGCGGCAAGAACCAAGTTGCGT 1800
 DB 2184 GACGGTGTCAAGAACTTAAACCGGCTTGTGTAGTGTGTGTGCAAGAACTTCCAAAGTTGAG 2243
 QY 1801 GAGTTGCTAACTCTGATGTTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
 DB 2244 GAATCGGCAACTTGT 2303
 QY 1861 GAGAAAGCTGAATGAGAAATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
 DB 2304 GAGCAGTCTGAGATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAG 2363
 QY 1921 TTCAGATGATATCTCTCAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAG 1980
 DB 2364 TTCAGATGATATCTCTCAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAG 2423
 QY 1981 TGCGACAGAAAGT 2040
 DB 2424 TGTGACAGAAAGT 2483
 QY 2041 GTGAGAGCAATGACTTGT 2100
 DB 2484 GTTGAAGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2543
 QY 2101 ATTATGT 2160
 DB 2544 ATCATTTGT 2603
 QY 2161 GACATATCTGT 2220
 DB 2604 GAGCTTCTGT 2663
 QY 2221 TCCCAAGAGGCTTGAAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAG 2280
 DB 2664 TCAAGAGGCTTGAAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAG 2723
 QY 2281 CTATTGACCTTGAAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAG 2340
 DB 2724 CTGTTGAACCTGAGCTGT 2783
 QY 2341 GAGAGT 2400
 DB 2784 GAGAGT 2843
 QY 2401 GTTCCATTGT 2422
 DB 2844 GTTCTCGGCTGT 2865
 RESULT 6
 ABZ13815
 ID ABZ13815 standard; DNA; 2427 BP.
 AC ABZ13815;
 XX
 DT 21-JAN-2003 (first entry)
 XX
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1620.
 XX
 KW Arabidopsis thaliana; plant; gene; stress; transgenic; de.
 XX
 OS Arabidopsis thaliana.
 XX
 PN MO20021655-A2.
 XX

PD 28-FEB-2002.
 PF 24-AUG-2001; 2001WO-US026685.
 XX
 XX 24-AUG-2000; 2000US-0227866P.
 PR 26-JAN-2001; 2001US-0264647P.
 PR 22-JUN-2001; 2001US-0300111P.
 XX
 PA (SCRI) SCRIPES RES INST.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Harper JF, Kreps J, Wang X, Zhu T;
 DR WPI; 2002-304127/34.
 XX
 PT Identifying a stress condition to which a plant cell has been exposed and
 PS producing plants with increased tolerance to these abiotic stresses.
 XX Claim 144; SEQ ID NO 1620; 577bp + Sequence Listing; English.
 CC The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array of probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office
 XX
 SQ Sequence 2427 BP; 643 A; 517 C; 587 G; 680 T; 0 U; 0 Other;
 Query Match 56.6%; Score 1486.6; DB 6; Length 2427;
 Best Local Similarity 76.3%; Pred. No. 0;
 Matches 1841; Conservative 0; Mismatches 570; Indels 3; Gaps 1;
 QY 4 GCTGAGGCTGCTCATCTGCGTGTCCACAGTCCCTGAGCGTTTGATGAGACCTTCTT 63
 DB 10 GCTAGCGTATGATAGACGCGGTCCACAGCCAGTGTGCTTTGAAACGAAACCTTGT 69
 QY 64 GCTCAGGAAGAGATTTTGGCTGTCTCAAGATCGAGGCAAGAAAGAAATT 123
 DB 70 TCTGAGGAAGAGATTTTGGCTGTCTCAAGATCGAGGCAAGAAAGAAATT 129
 QY 124 CTGCAACACATCAATTTATCTAGATTTGAAAGCTATCCCTGAAAGAAAGAAAG 183
 DB 130 TTACAGCAAAACAGATATCTGAAATGGAACCTTGTGCAAAACCGGAAGAA 189
 QY 184 CTGCTAATGTGTGATTTTGAAGATTTGAAGCTGTGCAAGAGCATGTGTGCT 243
 DB 190 CTGGAAGGTGTCTTCTTCTTGTGACCTTCTCAAACTCAGGAAGCAATGTGTGCA 249
 QY 244 CCATGGGTGACCTTGT 303
 DB 250 CCATGGGTGACCTTGT 309
 QY 304 GTTCAAGGCTTGT 363
 DB 310 CTCATGTCTTGT 369
 QY 364 CTGTGTGTGAGATTTCAATGAAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 423
 DB 370 CTGTGTGTGAGATTTCAATGAAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 429
 QY 424 TCATCATTTCCCGGCAACCTTTCAAAATCAATGATGATGATGATGATGATGATGATGATGAT 483
 DB 430 GCGTCTATCTTCTTCTTCAACCTCAGCAAAATGATGAAATGATGATGATGATGATGATGAT 489
 QY 484 GGTCACTTTTGGGAAATTTGTTCATGACAAAGAGAGATGACCTTTGCTGAAATTC 543

Db 490 CGTCAATTTATGGGCTTAAGCTCTTCCATGACAAAGAGAGTTTGCTTCCATGCTTAAGTTC 549
Qy 544 CTGAGAGTCCATTGTGCACAGGGCAAGAACTAGATGTTGAATGACAGAACTTCAAGAACTTG 603
Db 550 CTTGCTCTTCCAGGCCACACGGGCAAGAACTGATGTTGAGCGAGAAATTCAGAACTTC 609
Qy 604 AATGCTCTTCAACATGTTTGAAGAAAGCAGAGGATCTTTGGTAACCTTACCTTCCTGAG 663
Db 610 AACACTCTGCACACACCTTGAGAAACAGAAAGGATCTTACGACAGACTTAAAGTCCCAA 669
Qy 664 ACACCATGTGCCGATTTGGAACACCGGTTCCAGAAATCGGTTTGAAGAAGTGTGGGT 723
Db 670 ACACCTGTATGAAGATTTTGAAGCCCAAGTTTGAAGATTTGGTCTTGAAGAGGAGTGGGA 729
Qy 724 GACACCGAGAAACGGGTCTCGAGATGATCCAACTCTTTTGGATCTTCTTGAAGCAACT 783
Db 730 GACAAATGACAGACGGTCTCTTGAATGATACGTTCTTTTGGACCTTCTTGAAGCGGCTT 789
Qy 784 GATCCTTGCACCTTGAAGAGTCTTGGAGAAATCCCAATGATGTTCAATGTTGTGATT 843
Db 790 GATCCTTGCACCTTGAAGACTTTCTTGGAAAGATCCAAATGATGTTCAAGTTGTGATC 849
Qy 844 CTCACTCCCAACGAGATCTTGGCTCAAGCAATGTTTGGGGTATCCGACACCGGTGGG 903
Db 850 CTCTCTCCACATGTTTACTTGTCTCAGACCAATGTTCTTGGTTACCTGACACATGATGGA 909
Qy 904 CAGGTTGTTTACATCTTGGATCAAGTCCGAGCTTGAAGAAATGAGATGCTCTCCGTATA 963
Db 910 CAGGTTGTTTACATCTTGAATCAAGTCTGATCAAGTCTGATGATAGAGATGCTTCAAGTAT 969
Qy 964 AAGCAACAAGACTCAATCACTCCCTGGAATCCATATTAATTAATGATCTTCTGAT 1023
Db 970 AAGCAACAAGACTCAATCACTCCCTGGAATCCATATTAATTAATGATCTTCTGAT 1029
Qy 1024 GCTGTCGAGCAACATGCGGCTCAACGATTTGAGAAAGTATACGGAACAGACACTCGGAT 1083
Db 1030 GCGGTGAGAACTCACTGCGGTGAGACGCTCTCAGAGAGTTTATGATTTCTGATGATCTGAT 1089
Qy 1084 ATTTCTTGAGTACCTTTCAGAAACAGAAAGGAAATGTTTCGAAAATGAGATCTCAAGATT 1143
Db 1090 ATTTCTTGAGTACCTTTCAGAAACAGAAAGGAAATGTTTCGAAAATGAGATCTCAAGATT 1149
Qy 1144 GAAAAAGTCTGGCCCTACTTGGAAACCTTACAGAGAGATGTTGCTCATGAAATCTCCAAA 1203
Db 1150 G---AAGTCTGGCCCTATCTGAGACTTACACCGAGATGCTGCGGTGAGACTTACGAAA 1206
Qy 1204 GAGTTGACGAGACCGCAGATCTGATTCGAGAAACGAGGACCGCAATATGCTGACC 1263
Db 1207 GATTTGAATGGCAGACCTGACCTTATCTATGTTATCTGATGATGAGAAATCTTGTGCT 1266
Qy 1264 TCTTTGCTCGACATTAATTAAGTGTCAACAGTGCACCATGCGCCATGCTTTGGAGAG 1323
Db 1267 TCTTTATATGGCTCAAAAATTGGTGTCACTCAGTGTACATTTGCTCATGTCTTGAAGAAA 1326
Qy 1324 ACAAATAATCCAGATTCAGATTCATTTGGAAGAAAGCTTGAAGCAATATCAATTTGCTT 1383
Db 1327 ACAAAGTACCGGATTCGATATCTACCTGAGAAAGCTTGAAGCAATATCAATTTGCTCA 1386
Qy 1384 TGCCAAATTTACAGCTGATCTTTTTCAGATGAACCATACAGATTTTCATATCACAGTACT 1443
Db 1387 TGCCAGTTTCATGCGGATATTTTGGCAATGAACCACTGATTTTCATATCACTAGTACT 1446
Qy 1444 TTCCAGAAATTTGAGAGAGAGACACTGTTGTCAATAAGAGCCACACTGCTTTC 1503
Db 1447 TTCCAGAAATTTGCTGAGAGAGAGAAACTGTTGGGAGATTAAGAACCAACAGCCCTTT 1506
Qy 1504 ACTCTTCTGCTCTTACCGGTGTGTACATGATGATGATGTTTGAATCCAAATTTCAAC 1563
Db 1507 ACTCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1566
Qy 1564 ATTGTTTCCCTGGTGTGATATGAGATATATCTTCCCTTACACGAGAAAGAGAGGAGG 1623
Db 1567 ATTGTTTCCCTGGTGTGATATGAGATATATCTTCCCTTACACGAGAAAGAGGATGA 1626

Qy 1624 TTGAGCAATTTCCATCTCTGATGAGAAACCTTCTTTTACACCAAACTTGAATGAAGAA 1683
Db 1627 TTGACTAAGTTTCCATCTGATGATGAGAAACCTTCTTACACCAAACTTGAATGAAGAA 1686
Qy 1684 CACTTATGATGCTCAATGACCCGCAACAGCCAAATCTGTTTCAATGAGCCAGGCTTGT 1743
Db 1687 CACTTATGATGCTCAATGACCCGCAACAGCCAAATCTTCTTCAATGAGCCAGGCTTGT 1746
Qy 1744 GGTGTCAAGAACTTAAACCGACTCTGTGAGTGTGCGGCAAGAAACCCAAATGTCGTGAG 1803
Db 1747 GGTGTCAAGAACTTGTCAAGCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1806
Qy 1804 TTGGCTTAACTCTGATGTTGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAG 1863
Db 1807 CTAGCTTAACTTGTGTTGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1866
Qy 1864 AAGCTGAATAATGAAGAAATGTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1923
Db 1867 AAGCTGAATAATGAAGAAATGTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1926
Qy 1924 AGATGATATCATCTCAATGAACAGAAATCGAAATGTTGAATCTTACCGATACATTTGC 1983
Db 1927 AGGTGATCTCTCTCAGATGAGACCGGTTAAGAAACGTTGATGATGATGATGATGATGAT 1986
Qy 1984 GACACGAAGGTGCTTTGTGACAGCTGCAATTTGATGAGACCTTTGATGACAGTTGTG 2043
Db 1987 GACACGAAGGTGCTTTGTGACAGCTGCAATTTGATGAGACCTTTGATGATGATGATGATGAT 2046
Qy 2044 GAGGCAATGATCTTGGCTTGGCAACATTTGGCAACCTGTTAAGTGTGACAGCCGAGAT 2103
Db 2047 GAGGCAATGATCTTGGCTTGGCAACATTTGGCAACCTGTTAAGTGTGATGATGATGATGAT 2106
Qy 2104 ATTGTCCATGGAATCTGTTTCAACATTTGATCTTACCATGATGATGATGATGATGATGAT 2163
Db 2107 ATTGTCCATGGAATCTGTTTCAACATTTGATCTTACCATGATGATGATGATGATGATGAT 2166
Qy 2164 ATACTGTCGATTTCTTTGAAAAGTGAAGAAAGATCTGATCTGATGATGATGATGATGAT 2223
Db 2167 ACTTGTGATTTCTTCAACAGTGTGAGAGATCTGATCTGATGATGATGATGATGATGAT 2226
Qy 2224 CAAGGAGCTTGAAGCAATGAGAGAGATGATGATGATGATGATGATGATGATGATGAT 2283
Db 2227 AAAGAGGCTTCAAGAGATGAGAGAGAAATPACCTTGGCAAAATCTATTTCAACAGAGCTC 2286
Qy 2284 TTGACCTTGAACAGAGTGTGATGATCTGAGAGATTTTCAACCTTGAACCGCTGAG 2343
Db 2287 TTGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2346
Qy 2344 AGTGTGTTTACCTTGAAGATGTTTATGCTTAAAGTACCGTAACTGCTGAATCACT 2403
Db 2347 GCTGCGCTTACCTTGAAGATGTTTCTATGATGATGATGATGATGATGATGATGATGAT 2406
Qy 2404 CCAATGCGAGAGA 2417
Db 2407 CTTCTTGCACAGA 2420

RESULT 7
ADA68333
ID ADA68333 standard; DNA; 2427 BP.
XX ADA68333;
XX AC
XX 20-NOV-2003 (first entry)
XX DE
XX Arabidopsis thaliana gene, SEQ ID 808.
XX DE
XX Plant; Bacterial infection; fungal infection; viral infection; ds.
XX Arabidopsis thaliana.
XX OS
XX
XX W02003000898-A1.

XX 03-JAN-2003.
PD 22-JUN-2001; 2001WO-1B001105.
XX 22-JUN-2001; 2001WO-1B001105.
XX 22-JUN-2001; 2001WO-1B001105.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX Claim 6; SEQ ID NO 808; 899bp; English.
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
SQ Sequence 2427 BP; 643 A; 517 C; 587 G; 680 T; 0 U; 0 Other;
Query Match 56.6%; Score 1486.6; DB 7; Length 2427;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 1841; Conservative 0; Mismatches 570; Indels 3; Gaps 1;
QY 4 GCTGAGCGTCTCTCAGTCGCGTCCAGCTCTCCGTGAGCGTTGGATGAGACCTTTCTT 63
DB 10 GCTGAGCGTCTCTCAGTCGCGTCCAGCTCTCCGTGAGCGTTGGATGAGACCTTTCTT 69
QY 64 GCTGAGCGTCTCTCAGTCGCGTCCAGCTCTCCGTGAGCGTTGGATGAGACCTTTCTT 123
DB 70 TCTGAGAGAGAGAGAGTCTCTGCTTCTTCCAGGCTTGAAGCCAGAGAGAGTATT 129
QY 124 CTGCAACCAATCAATATTATCTAGAGTTGAGAGTATCTCTGAGAGAGAGAGAG 183
DB 130 TTACAGAGAGAGAGAGTCTCTGCTTCTTCCAGGCTTGAAGCCAGAGAGAGTATT 189
QY 184 CTCGTAATGTCATTTTGAAGTATTGAAGGCTAGTCAAGAGAGAGTGTGCTT 243
DB 190 CTGAGAGAGAGAGAGTCTCTGCTTCTTCCAGGCTTGAAGCCAGAGAGAGTGTGCTT 249
QY 244 CCATGAGAGAGAGAGTCTCTGCTTCTTCCAGGCTTGAAGCCAGAGAGAGTGTGCTT 303
DB 250 CCATGAGAGAGAGAGTCTCTGCTTCTTCCAGGCTTGAAGCCAGAGAGAGTGTGCTT 309
QY 304 GTTCAAGCGCTTGTGTTGAGAGAGTCACTGTTGCTGATCTTCCATTCAGAGAGAG 363
DB 310 CTCATAGCTCTTGTGTTGAGAGAGTCACTGTTGCTGATCTTCCATTCAGAGAGAG 369
QY 364 CTTGTTGAGAGAGTCAATGAGAGAGTGTGTTGAGAGAGTGTGAGAGAGAGAG 423
DB 370 CTCGTTGAGAGAGTCAATGAGAGAGTGTGTTGAGAGAGTGTGAGAGAGAGAG 429
QY 424 TCATCATTCCTCCGCGCAACTCTTCAAAATCATTTGAGAGAGAGAGAGAGAG 483
DB 430 GCGTATATCCCTGCGTCAACATCAAAATCATTTGAGAGAGAGAGAGAGAG 489
QY 484 GGTCACTTTGCGCAAAATTTGTTCCATGACAGAGAGAGAGAGAGAGAGAGAG 543
DB 490 GGTCACTTTGCGCAAAATTTGTTCCATGACAGAGAGAGAGAGAGAGAGAGAG 549

QY 544 CTCAGATTCATTTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 603
DB 550 CTTGCTCTTCAAG 609
QY 604 AATGCTCTTCAAG 663
DB 610 AATGCTCTTCAAG 669
QY 664 ACACAG 723
DB 670 ACACAG 729
QY 724 GACACGAG 783
DB 730 GACACGAG 789
QY 784 GATCCTTCAAG 843
DB 790 GATCCTTCAAG 849
QY 844 CTCACCTCCAG 903
DB 850 CTCACCTCCAG 909
QY 904 CAGGTTGTTACATCTTGAGATCAAGTCGAGCTTGGAGAGAGAGAGAGAGAGAGAG 963
DB 910 CAGGTTGTTACATCTTGAGATCAAGTCGAGCTTGGAGAGAGAGAGAGAGAGAGAG 969
QY 964 AAGCAAG 1023
DB 970 AAGCAAG 1029
QY 1024 GCTGTGAG 1083
DB 1030 GCTGTGAG 1089
QY 1084 ATTCTTGAG 1143
DB 1090 ATTCTTGAG 1149
QY 1144 GAAAAAGCTGAG 1203
DB 1150 G---AAGTGGGAG 1206
QY 1204 GAGTTGAG 1263
DB 1207 GAAATTGAG 1266
QY 1264 TCTTGTGAG 1323
DB 1267 TCTTGTGAG 1326
QY 1324 ACAAAATATCCAGATTCAGATTCATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1383
DB 1327 ACAAAATATCCAGATTCAGATTCATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1386
QY 1384 TGCAGATTCAGATTCATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1443
DB 1387 TGCAGATTCAGATTCATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446
QY 1444 TTTCAAG 1503
DB 1447 TTTCAAG 1506
QY 1504 ACTCTTCCTGCTCTCAAG 1563
DB 1507 ACTCTTCCTGCTCTCAAG 1566
QY 1564 ATTGTTTCCCTGCTCTCAAG 1623
DB 1567 ATTGTTTCCCTGCTCTCAAG 1626

D	b		471	CGTCTTCATGTCGAAATGTCATTCGAAATGGGGTGCAGTTCTTAAACGCTCACTGCT	530
O	y		496	GCAAAATGTTTCATGACAGAGAGCATGCAACCTTTGCTCGAATTCCTCAAGTCCAT	555
D	b		531	TCCAGCTGTTTCAGAGACAGAGAGAGCTCTACCCACTACTGAATCTTCTGAAGCCCAT	590
O	y		556	TGTCAAGAGGGCAAGACATGATGTGTGAATGACAAATTCAGAACTTGAATGCTTCTTCA	615
D	b		591	AACCAAGAGGGACACATGATGTCTAAACGACCGAATTCAGAGCTTCGTGGGCTCCAA	650
O	y		616	CATGTTTGAAGAAAGCAGAGAGATCTTGATCCCTACCTCCAGAACACATGTCGC	675
D	b		651	TCAGCCCTTAAGAAAGCAGAGAGATCTTAACGACATTCCTGAAGACACCCCGTCTCT	710
O	y		676	GAATTCGAAACCCGGTTCAGGAAATCGGTTTGAAGAAGGTTGGGGTGAACCCGACAA	735
D	b		711	GAGTTCACACACAGGTTCCAAAGAGCTTGAGTGGAGAAAGGGTTGGGGCGACACGCTAAG	770
O	y		736	CGCGTCTCGAGATGATCTCACTCTTTTGGATCTTTTGAAGGCACTGATCTTTGACC	795
D	b		771	CGTGAACGAGACACCATCATCTTGCTTCTTGAATCTTAAGAGGCCCTCGATTCACGACGC	830
O	y		796	CTTGAAGAGTTCCTTGGGAGAAATCCCAATGGGTCAATGTTGTGATCTCTACCTCCAC	855
D	b		831	TTGGAGAAAGTTCCTTGGAACTTACCAATGATGTTCAATGTTGTTATCTGTCTTCACAC	890
O	y		856	GGATATCTTCGCTCAAGACATGTTTGGGGTATCCCGAACCCGGTGGGCAAGTGTGTTAC	915
D	b		891	GGATATTTTGGCTCAATCCAAATGTTGGGATACCTGTAACTGATGGCCAGGTTGTGATC	950
O	y		916	ATCTTGGATCAAGTCCGAGCTTTGGAGATGAGATGCTCTCCGTATTAAGCAACAGAA	975
D	b		951	ATCTTGGATCAAGTCCGCTTGGAGATGAGATGCTCTGAGGATTAAGACAGACAGCC	1010
O	y		976	CTCAACATCACCCCTCGAATCTCTATTTACTACATCTTCCGATGCTGTGGAGCA	1035
D	b		1011	CTTGACATTAACCTTAAGTCTCAATGTCACACAGGCTGTGCTCGATGCTGTGGAAT	1070
O	y		1036	ACATGCGGTCAACGACTTGAGAAATATACGGAACAGACACTCGAATATTCCTCGAGTA	1095
D	b		1071	ACATGTGGCCAGCGGCTGGAGAAATTAATTTGGGACTGAGACACTGACATTCCTCGGT	1130
O	y		1096	CCCTTCAGAACAGAAAGGAAATTTGTGAAATGAGATCTCAAGATTTGAAAAAGTCTGG	1155
D	b		1131	CCATTCAAGACTGAGAAAGGGGACTCTCGTAAGTGAATCTCTCGTTTG---ATGTCTGG	1187
O	y		1156	CCATPCTTGGAAACCTTACACAGAGATGTTGTCTATGAAATCTCCAAAGATGTGACGGC	1215
D	b		1188	CCATPCTTGGAGACATPACCGAGGATGTGAAACGAATCATGAGAGAAATGTGAGACC	1247
O	y		1216	ACGCCAGATCTGATCATCGSAAACAAACAGCGACGCGCAATATCGTGCCTCTTGCTCGCA	1275
D	b		1248	AAGCTGATCTCATTAATTTGCGACCTACAGTGAACGCTTAACCTTGTGCGACTCTGTGGG	1307
O	y		1276	CATAAATTAAGTGTCAACAGTGCAACATCGCCCAATGCTTTGGAGAGACAAATATATCA	1335
D	b		1308	CATAAAGTTGGAGTTAACCCAGTGCACTTGGCCCATGCTTTGGAGAAAAACAAATATACCC	1367
O	y		1336	GATTCAGATCTTAATTTGGAAGAAGCTTGAAGACAAATACATTTCTCTTGGCAATTTACA	1395
D	b		1368	AACCTGACATTAATTTGGAACAAATTTGACACGCAATACATTTCTCAATGCAAGTCACT	1427
O	y		1396	GCTGATCTTTTGGCAATGACACATCAATTTCAATCACACAGTACATTTCCAGGAAT	1455
D	b		1428	GCTGACCTTAATTTGCCATGATACACTGATTTCACTACACAGACATTTCCAGGAAT	1487
O	y		1456	GCAAGAGCAAGACACTGTTGGTCAATACAGAGCCCACTGCTTCACTCTTCTGGT	1515
D	b		1488	GCTGGAACCAAGACAGCGTGGGCGAGATGATGCTCACTTGGCTTCACTCTCTGAT	1547
O	y		1516	CTTACCGTGTGATCATGTATCGATGATGTTTGAATCCCAATTCACATTTGTTCCCT	1575

Db	1548	CTTACCGTGTGTCACATGGATATGACGTTGATCCCAAGTTCAACATTTGCTTCCT	1607
Oy	1548	CTTACCGTGTGTCACATGGATATGACGTTGATCCCAAGTTCAACATTTGCTTCCT	1607
Db	1576	GGTGTGATATGAGGATATATCTTCCCTTACACCGAAGAAAGCGAGTTGAACATTTCC	1635
Oy	1576	GGTGTGATATGAGGATATATCTTCCCTTACACCGAAGAAAGCGAGTTGAACATTTCC	1635
Db	1608	GGACGAGATATGACTGTCTCACTTCCATPACCTGAACCTGACAAAGAGCTTACGCCCTTC	1667
Oy	1608	GGACGAGATATGACTGTCTCACTTCCATPACCTGAACCTGACAAAGAGCTTACGCCCTTC	1667
Db	1636	CATCTTGAGATCGAAGACCTTCTTTACACCAAGTTGAGATGAAGAACTATATGTGTG	1695
Oy	1636	CATCTTGAGATCGAAGACCTTCTTTACACCAAGTTGAGATGAAGAACTATATGTGTG	1695
Db	1668	CACCTTGAAATGGAAGAGCTCTTTACAGTATGATTTGAGAACTTGAAACACAAATTTGTA	1727
Oy	1668	CACCTTGAAATGGAAGAGCTCTTTACAGTATGATTTGAGAACTTGAAACACAAATTTGTA	1727
Db	1696	CTCAATGACCGCAACAAGCCAAATTCGTTCACAAATGCGCAAGGCTTGATCGTGCAGAAC	1755
Oy	1696	CTCAATGACCGCAACAAGCCAAATTCGTTCACAAATGCGCAAGGCTTGATCGTGCAGAAC	1755
Db	1728	TTGAAGGACAAAGAACAGGCTATCATCTTCCATATGCTGTGCTTTGACCGGTGAAGAAC	1787
Oy	1728	TTGAAGGACAAAGAACAGGCTATCATCTTCCATATGCTGTGCTTTGACCGGTGAAGAAC	1787
Db	1756	TTAAACCGGACTCGTCGAGTGTGCGGCAAGAACCCAAAGTTGCGTGAATTGCACTTC	1815
Oy	1756	TTAAACCGGACTCGTCGAGTGTGCGGCAAGAACCCAAAGTTGCGTGAATTGCACTTC	1815
Db	1788	ATGACAGGCTTGTTGAGATGTTTGGTAAAGATCTCATCTGAAGAGATTTGGCAAACTTT	1847
Oy	1788	ATGACAGGCTTGTTGAGATGTTTGGTAAAGATCTCATCTGAAGAGATTTGGCAAACTTT	1847
Db	1816	GTACTTTGAGTGTGATATGGCGAAAGAACTPAAATTTGGAAGAAAGGCTGAAGATG	1875
Oy	1816	GTACTTTGAGTGTGATATGGCGAAAGAACTPAAATTTGGAAGAAAGGCTGAAGATG	1875
Db	1848	GTGATGTGCTGTGTGACCATGCGAAGAGTCAAGATATGGAGGAGCAAGGCTGAAGTTC	1907
Oy	1848	GTGATGTGCTGTGTGACCATGCGAAGAGTCAAGATATGGAGGAGCAAGGCTGAAGTTC	1907
Db	1876	AAGAAATATGTTTGAAGCTGATCGACAAATPACAACTTGAACGGCCAAATTCAGATGATATCA	1935
Oy	1876	AAGAAATATGTTTGAAGCTGATCGACAAATPACAACTTGAACGGCCAAATTCAGATGATATCA	1935
Db	1908	AAAGAGATGATCACTTCATTTGAGAGATCAAGCTGAGGCGCAATTCGGTGGATCTCA	1967
Oy	1908	AAAGAGATGATCACTTCATTTGAGAGATCAAGCTGAGGCGCAATTCGGTGGATCTCA	1967
Db	1936	TCTCAATATGAACAGAAATCCGAATGTTGAACCTTATCCGATACATTTGCGACAGAAAGT	1995
Oy	1936	TCTCAATATGAACAGAAATCCGAATGTTGAACCTTATCCGATACATTTGCGACAGAAAGT	1995
Db	1968	GCTCAGATATGAACCGTGTTCGCAATGCAAGAGTGTACCGCTPACATTTGTGACACCAAGGGC	2027
Oy	1968	GCTCAGATATGAACCGTGTTCGCAATGCAAGAGTGTACCGCTPACATTTGTGACACCAAGGGC	2027
Db	1996	GCCTTTGTACAGCCTGATTTGTATGAAGCCTTTGGAATTTGACATTTGTGTGAGGCAATGACT	2055
Oy	1996	GCCTTTGTACAGCCTGATTTGTATGAAGCCTTTGGAATTTGACATTTGTGTGAGGCAATGACT	2055
Db	2028	GCAATTTGTTCAGCTGTGATCTTATGAAGCAATTTGGCTGTGCTGTGTGCGAGGCATAGCA	2087
Oy	2028	GCAATTTGTTCAGCTGTGATCTTATGAAGCAATTTGGCTGTGCTGTGTGCGAGGCATAGCA	2087
Db	2056	TGCGATTTGCGCAATTCGCAACTGTATACGCTGACCAAGCGAGATTAATTTGTCATGAGG	2115
Oy	2056	TGCGATTTGCGCAATTCGCAACTGTATACGCTGACCAAGCGAGATTAATTTGTCATGAGG	2115
Db	2088	TGTGTGTTTCCCAATATGAGCAATGCAATGCAATGTGTGCGCAAGCTGAATTCATTTGTGAATGCT	2147
Oy	2088	TGTGTGTTTCCCAATATGAGCAATGCAATGCAATGTGTGCGCAAGCTGAATTCATTTGTGAATGCT	2147
Db	2116	AAATCTGGTTCACATTTGATCTTTACCAATGATGATCAAGCTGTGACATATCTGTCGAT	2175
Oy	2116	AAATCTGGTTCACATTTGATCTTTACCAATGATGATCAAGCTGTGACATATCTGTCGAT	2175
Db	2148	GTTGCTGGTTCACATTCGATCTTTACCAAGTGAACAAGGCTGCAATATCTGTCGATCAAC	2207
Oy	2148	GTTGCTGGTTCACATTCGATCTTTACCAAGTGAACAAGGCTGCAATATCTGTCGATCAAC	2207
Db	2176	TTCTTTGAAAAAGTAAAGAAATCCATCTCACCTGGATTAAGATCTCCCAAGAGGCTTG	2235
Oy	2176	TTCTTTGAAAAAGTAAAGAAATCCATCTCACCTGGATTAAGATCTCCCAAGAGGCTTG	2235
Db	2208	TTCTTTGAGAAATCAGCGGACGATCCAACTTACGGGACAAAAAGTGTGGAAGGTGGCGCTG	2267
Oy	2208	TTCTTTGAGAAATCAGCGGACGATCCAACTTACGGGACAAAAAGTGTGGAAGGTGGCGCTG	2267
Db	2236	AAACGATATGAGGAGAGATATACATGAGAGATTTACTCGAGAGACTATTTGACCCCTGACA	2295
Oy	2236	AAACGATATGAGGAGAGATATACATGAGAGATTTACTCGAGAGACTATTTGACCCCTGACA	2295
Db	2268	AAGGAAATTTATGAGAAAGTACACTGGAAGCTGTATCTCAAGAGAGGCTGATGACCTTGACT	2327
Oy	2268	AAGGAAATTTATGAGAAAGTACACTGGAAGCTGTATCTCAAGAGAGGCTGATGACCTTGACT	2327
Db	2296	GGAAGTATATGATCTCTGGAAGCAATGTTTCCAACTTTGAAAGCGCCGCTGAGAGTCTGCTTAC	2355
Oy	2296	GGAAGTATATGATCTCTGGAAGCAATGTTTCCAACTTTGAAAGCGCCGCTGAGAGTCTGCTTAC	2355
Db	2328	GGTGTATATGTTTTCTGGAAGTATGTGAGCAACTTTGAGAGGCGTGAAGCTGCGCGTTAC	2387
Oy	2328	GGTGTATATGTTTTCTGGAAGTATGTGAGCAACTTTGAGAGGCGTGAAGCTGCGCGTTAC	2387
Db	2356	CTTGAGATGTTTATGCTCTTAAAGTACCGTAAAGCTGCTGATATATGTTCCATTGGCAGAG	2415
Oy	2356	CTTGAGATGTTTATGCTCTTAAAGTACCGTAAAGCTGCTGATATATGTTCCATTGGCAGAG	2415
Db	2388	CTTGAGATGTTTCTACGCTCTGAAAGTACCGGACGCTGCTGCTGCAAGTTCCATTTGCCGCTC	2447
Oy	2388	CTTGAGATGTTTCTACGCTCTGAAAGTACCGGACGCTGCTGCTGCAAGTTCCATTTGCCGCTC	2447
Db	2416	GA 2417	
Oy	2416	GA 2417	
Db	2448	GA 2449	
Oy	2448	GA 2449	
RESULT 9			
ID	ADDC68520	standard; cDNA; 2716 BP.	
XX	ADDC68520;		
XX	AC		
XX	18-DEC-2003	(first entry)	
XX	DT		
DE	Lolium perenne fructan biosynthesis protein cDNA SEQ ID NO:12.		
XX			

KW *Lolium perenne*; *Festuca arundinacea*; lignin; fructan; tannin; biosynthetic pathway; plant; gene; ss.

OS *Lolium perenne*.

PN WO2003040306-A2.

PD 15-MAY-2003

PF 07-NOV-2002; 2002WO-NZ0000239.

PR 07-NOV-2001; 2001US-0337703P.

PA (GENE-) GENESIS RES & DEV CORP LTD.

XX

PI Saulsbury KM, Hall C;

DR WPI; 2003-441544/41.

XX

PT arundinacea, useful for modulating the biosynthesis of lignin, fructan or

XX

The present invention describes isolated polynucleotides (I) encoding proteins (II) from *Lolium perenne* and *Festuca arundinacea* which are active in lignin, fructan and tannin biosynthetic pathways. Also described: (1) an isolated oligonucleotide probe or primer comprising at least 10 contiguous residues complementary to 10 contiguous residues of (1); (2) a kit comprising the oligonucleotide probe or primer; (3) a genetic construct comprising (1); (4) a transgenic plant cell comprising the genetic construct of (3); (5) a plant or its seed, fruit or progeny comprising the transgenic plant cell of (4); (6) modulating one or more of the lignin, fructan or tannin compositions of a plant; (7) producing a plant having one or more of the lignin, fructan or tannin compositions; and (8) modifying the activity of (II) involved in a lignin, fructan or tannin biosynthetic pathway in a plant. (I) can be used for modulating the biosynthesis of lignin, fructan or tannin in a plant. The present sequence is used in the exemplification of the present invention.

SQ Sequence 2716 BP; 656 A; 666 C; 672 G; 722 T; 0 U; 0 Other;

Query Match 49.78; Score 1305.4; DB 9; Length 2716;

Matches 1731; Conservative 0; Mismatches 662; Indels 9; Gaps 2;

Qy	76	GAGATTTTGGCCCTTGCCTCAGAGATCGAGGGCAAAAGAAAAGAAATTCGCAACACAT	135
Db	58	CTGACCTGCCTCACACAGTCTCCGGGAGGCCCTTGGTGCACCTTCTCCCTCACTTAAC	117
Qy	16	CTCACTGCCTGCACACAGTCTCCGGTGGCGGTTTGAATGAGACCCCTCTTGCACAGGAGC	75
Dy	118	GAGCTCATTTCCCTCTTTTCAAAAGTATGTTTCAACAGGGGAAAGAGATCTTCAGCGCAC	177
Qy	136	CAATATTTCAGAGTTTGAAGCTATCCCTGGAAGAAACAGAAAGAACTCGTAAATGCT	195
Dy	178	CAGGTGCTCATGATGTTTGAAGCTTGTTTGAAGGCTGACAAAGGAGA-----GATTTAGCA	231
Qy	196	GCATTTTGTGAAGTATTGAAGGCTAGTCAGGAAGGAGATCTGTGTGCTTCATGCGTTGCA	255
Dy	232	CCCTTTGAAAGACATTTCTCCGTGTGCTCAGGAAGCAATTTGCTCCGCCCAATGGTTTGA	291
Qy	256	CTTGCTGTTGTCCTCAAGGCTGTGTGTTTGGAGTACATTAGAGTAATGTTCAAGCCCTT	315
Dy	292	CTGGCCATTCAGGCTTGAAGCTGTGTCTGGACATCATACGGGTGAAATGTTAGGAATTG	351
Qy	316	GTTGTGTAGGAATCACTGTTGCTGAGTATCTCCACTTCAGGAAGAACTGTTGTATGGA	375
Dy	352	GCCTTTGAGGAGCTGACAGTTTCCGAGTACCTGGCAATTCAGGAACACCTGTTGATGAG	411

QY	376	AGTTCAATGAAACTTTGTTTGGATTTGGATTTTGAAGCCCTTCACTCATCTATTTCCC	435
Db	412	CAGCCGACAGCAAGTTTGCTGGACCTTGATTTTGAAGCTTTCATATGCTCTTCCCA	471
QY	436	CGCCCACTCTTTCAAAATCCATTTGGTAATAGTGTGAGTTCTTAATTCGATCCTTTG	495
Db	472	CGTCTTTCATGTCCAAAGTCATTGGAAATGGGGTGAGTTTCTTAAACCGTCACTGTCT	531
QY	496	GCAAAATTTGTTCCATGACAAAGAGAGCATGACCCCTTTGCTCGAATTCCTCAGAGTCAT	555
Db	532	TTCAAGCTGTTCCAGACAAAGAGAGGCTCTACCACTACTGAACTTCTCGAAGGCCAT	591
QY	556	TGTGACAAAGGCGCAAAACATGATTTGAAATGACAAATTCGAACCTTGAATGCTTTCAA	615
Db	592	AACCAACAGGGCACGAAATGATGCTTAAAGCAGCCGAATTCAGAGCCTTCGGGGCTTCAA	651
QY	616	CATGTTTTGAGGAAGAGAGAGTATCTTGATCCCTACCTCCTGAGACCATGAGCC	675
Db	652	TCAGCCCTTGAAGAAAGCAGAAAGATATCTAACGAGCATCTCTGAAGACACCCTCTCT	711
QY	676	GAATTCGACACCGGTTCCAGAAATCGTTTGGAAAGAGTTGGGGTGACACCGCAGAA	735
Db	712	GAGTTCAACCAAGGTTTCCAAAGGCTTGCTTGGAGAAAGGTTGGGGCGACACCGCTAAG	771
QY	736	CGCGTGTCTGAGATGATCCACTCTCTTTGGATCTTTGAGCACTGATCCTTGACCC	795
Db	772	CGTGTAACGAGCACCATTCACCTTGCTCTTGATCTTACCTTGAGGGCCCTGTATCCAGCAGC	831
QY	796	CTTGAGAAAGTTCCTTGGGAGAAATCCCATAGTGTTCAAATGTGAGATTTCACTGCCAC	855
Db	832	TTGAGAAAGTTCCTTGGAACTTATCAATATGTTCAATGTGTTATCTGTCTCCACAC	891
QY	856	GGATACTTTCGCTCAAGCAATGTTTGGGGTATCCGACACCGGTGGCAGGTGTTTAC	915
Db	892	GGATATTTTGTCTCAATCCATGTGTGGGATACCTCTATCTGTGGCCAGGTTGTGTAC	951
QY	916	ATCTTGGATCAAGTCCGAGCTTTGGAAATGAGATGCTCTCCGTATPAAAGCAACAAGA	975
Db	952	ATCTTGGATCAAGTCCGAGCTTTGGAAATGAGATGCTCTTGAGGATTAAGACACAGGC	1011
QY	976	CTCAACATCACCCCTCGAATCTCATTTATTACTTGATCTTCTCTGATGCTGTGGAACA	1035
Db	1012	CTTGACATPAAACCCCTTAAGATCTCATTTGTCAACCGGTGTTGCCGTGATGCTGTGAACT	1071
QY	1036	ACATGCGGTCAACACTTTGAGAAAGTTAACGGAACAGAGCACTCGAATATCTTCAGATA	1095
Db	1072	ACATGTGGCCAGCGGCTGGAGAAAGTTATTTGGGACTGAGCACTGACATTTCTCCGTGT	1131
QY	1096	CCCTTTCAGAACAAAAAGGAAATGTTTCGAAATGATCTCAAGATTTGAAAAAGTGTGG	1155
Db	1132	CCATTTGAAATGAGAAAGGGGATCTCCGTAATGAGATCTCTCGTTTG---ANGTCTGG	1188
QY	1156	CCATATCTTGGAAACCTTACACAGAGAGTGTGCTATGAAATCTCCAAAGATTGCAAGGC	1215
Db	1189	CCATATCTTGGAGATCACCGAGAGATTTGCAAAACCAATCAGAGAAATGCAAGCC	1248
QY	1216	AACGCAATCTGATCATCTGAGAAACAAAGGAGCGGCAATATCGTCGCTCGTGTGCA	1275
Db	1249	AAGCTGATCTCATTTATTTGGCACTTACAGTACGGGTAACTTTGTGCGCACTCTGCTTGGC	1308
QY	1276	CATPAAATTTAGGTTCACACAGTGCACCATCGCCCATGCTTTGAGAGAACAAATATCA	1335
Db	1309	CATPAGTTGGAGATTACCGAGTGCACATTTGGCCATGCTTTGGAGAAAAAATAATACCC	1368
QY	1336	GATTCAATATCTAATTGGAAGAGCTTGAAGAACAAATACCAATTTCTTGTGCAATTTACA	1395
Db	1369	AACTCAGACATATTTTGGCAAAATTTGACAGCGCAATATCAATTTCTCATGCAAGTTACT	1428
QY	1396	GCATATCTTTTGGAAAGAACATACAGATTTCACTATCAACGATCTTTCCAGGAATT	1455
Db	1429	GCTACCTTTATGGCAATGATCACTGATTTTCACTATCAACGACATTTCCAGGAATT	1488

QY 1456 GCAGGACGAGACGCTGGTGCATATACAGAGCCACACTGCTTTCATCTTCCTGAT 1515
 DB 1489 GCTGGACACAGAGGATGGGCGAGATGATGATGATGATGATGATGATGATGATGAT 1548
 QY 1516 CTCTACCGTGTGATACATGATGATGATGATGATGATGATGATGATGATGATGAT 1575
 DB 1549 CTCTACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1608
 QY 1576 GGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1635
 DB 1609 GAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1668
 QY 1636 CATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1695
 DB 1669 CACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1728
 QY 1696 CTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1755
 DB 1729 TTGAAGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1788
 QY 1756 TTAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1815
 DB 1789 ATGACAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1848
 QY 1816 GTAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1875
 DB 1849 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1908
 QY 1876 AAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1935
 DB 1909 AAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1968
 QY 1936 TCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1995
 DB 1969 GCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2028
 QY 1996 GCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2055
 DB 2029 GCATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2088
 QY 2056 TGCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2115
 DB 2089 TGTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2148
 QY 2116 AAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2175
 DB 2149 GTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2208
 QY 2176 TTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2235
 DB 2209 TTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2268
 QY 2236 AAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2295
 DB 2269 AAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2328
 QY 2296 GAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2355
 DB 2329 GGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2388
 QY 2356 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2415
 DB 2389 CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2448
 QY 2416 GA 2417
 DB 2449 GA 2450

RESULT 10
 ID ABR98516 standard; cDNA; 2829 BP.
 XX

AC ABR98516;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE cDNA sequence encoding perennial ryegrass sucrose synthase.
 XX
 DB Perennial ryegrass; sucrose phosphatase synthase; SPS; invertase; INV;
 XX sucrose synthase; SS; sucrose:sucrose 1-fructosyltransferase; SST;
 KW sucrose transporter; ST; fructan:fructan 1-fructosyltransferase; PFT;
 KW sucrose:fructan 6-fructosyltransferase; SFT; soluble carbohydrate transport;
 KW SFT; soluble carbohydrate metabolism; plant; gene; ss.
 XX
 OS Lolium perenne.
 XX
 XX WO20021130-A1.
 XX
 PD 18-APR-2002.
 XX
 PF 10-OCT-2001; 2001WO-AU001275.
 XX
 PR 10-OCT-2000; 2000AU-00000673.
 XX
 PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
 XX (AGRE-) AGRESEARCH LTD.
 PI Spangenberg G, Sawbridge TI, Ong EK, Emmerling M;
 XX WPI: 2002-444177/47.
 DR P-PSDB; ABR69063.
 XX
 PT New enzymes and transporter proteins involved in metabolism and/or
 PT transport of soluble carbohydrates, useful as molecular genetic markers,
 PT and in modifying soluble carbohydrate metabolism and/or transport in
 PT plants.
 XX
 PS Claim 5; Fig 74; 177bp; English.
 XX
 CC The present invention relates to the isolation of polynucleotide
 CC sequences encoding perennial ryegrass enzymes or transporter proteins
 CC selected from sucrose phosphatase synthase (SPS), invertase (INV), sucrose
 CC synthase (SS), sucrose transporter (ST), sucrose:sucrose 1-
 CC fructosyltransferase (SFT), fructan:fructan 1-fructosyltransferase (PFT),
 CC and sucrose:fructan 6-fructosyltransferase (SFT). The nucleic acid
 CC sequences of the invention are useful as molecular genetic markers, to
 CC isolate cDNAs and genes encoding homologous proteins from the same or
 CC other plant species, and for the modification of soluble carbohydrate
 CC metabolism and/or transport in plants. The polypeptide sequences may be
 CC used to immunise animals to produce antibodies with specificity for the
 CC proteins, and these antibodies may then be used to screen cDNA expression
 CC libraries to isolate full-length cDNA clones of interest. ABR98425-
 CC ABR98516 represent nucleotide sequences encoding perennial ryegrass
 CC enzymes or transporter proteins involved in the metabolism and/or
 CC transport of soluble carbohydrates
 XX
 SQ Sequence 2829 BP; 697 A; 678 C; 697 G; 757 T; 0 U; 0 Other;

Query Match 49.7%; Score 1305.4; DB 6; Length 2829;
 Best Local Similarity 72.1%; Pred. No. 0;
 Matches 1731; Conservative 0; Mismatches 662; Indels 9; Gaps 2;

QY 16 CTCACGCGCGCCACAGCTCCGATGAGGCTTGTGATGAGACCTTCTTGTGCACAGAAC 75
 DB 111 CTGACCTGCTCCACAGCTCTCCGAGCGCTTGTGATGAGACCTTCTTGTGCACAGAAC 170
 QY 76 GAGATTTTGGCTTCTCTCAAGATGAGGCGCAAGGAAAGAAATTTGCAACACCAT 135
 DB 171 GAGCTCATTTGCTCTTTTCAAGATGATGATGATGATGATGATGATGATGATGAT 230
 QY 136 CAATATATCTAGAGTTGAAAGCTATCCCTGAAGAAGAAAGAAAGCTGCTAATGCT 195
 DB 231 CAGCTGCTCAGTGAAGTTGAAAGCTTTGTTGAGGCTGACAAAGAGA-----GATATGCA 284
 QY 196 GCATTTTGAAGTATGAAGCTAGTCAGAGACGATGCTGTTGCTCATGAGGTTGCA 255

```
Db 285 CCGTTGAAGACATTAATCGGTGCTGACGAGAGAAATGTGTGCCCCACAGGGTGTGCA 344
Qy 256 CTTGTGCTTCGTCGAGAGCGCTGTGTGTTGGAGATACATTAGATGTAATGTTCAAGCCCTT 315
Db 345 CTGGGCATCAGAGCCGAGCCGAGTGTCTGGGACATACACGGGTGAATGTTAGGAGAGT 404
Qy 316 GTTGTGAGGAATCTCATGTGTGCTGAGTATCTCACTTCAAGAAAGACTTGTGTAGTA 375
Db 405 GCTGTGAGGAGCTGACAGTTTCCAGATACCTAGACATCAAGAAACAGCTTGTGTAGAG 464
Qy 376 AGTTCAATGGAACCTTGTGTTGGAAATTTGAGCCCTTCAACTCATCAATCTCC 435
Db 465 CACGACACACAGATTTGTGCTTGAAGCTTAATTTGAGCCCTTCAATGCTTCTCCG 524
Qy 436 CGCCCAACTCTTTCAAAATTCATTGTAATGTTGAGTTCTTAATCTCACTTCACTTTCG 495
Db 525 GGTCTTCATGTCGAAATCTCATTTGGAATGGGGTGCACTTCTTAACCGTCACTGTCT 584
Qy 496 GCAAAATTTGTTCCATGACAGAGAGACATGCAACCTTGTCTGAAATTCCTGAGATCCAT 555
Db 585 TCCAAGCTGTTCCAGGACAGGAGAGCCTTAACCACTAGAACTTCTGAAAGCCCAT 644
Qy 556 TGTCAAGAGGCGAAGAAATGATGTTGAATGACAGAAATTCAGAACTTGAATGCTCTTCAA 615
Db 645 AACTACAAAGGACACGACAAATGCTAAATGACAGAAATTCAGAGCTTGTGAGGCTCCAG 704
Qy 616 CATGTTTGAAGAAAGACAGAGAGATCTTGGTACCTTACCTCTGAGACACATGTC 675
Db 705 TCAGCCCTTAAGAAAGCAGAAAGATTTTAAGAGACATCCCTGAGAGACACCCGCTCCT 764
Qy 676 GAATTCGAAACCGGTTCCAGAAATCGTTTGAAGAGGTTGGAGTGAACCGCAGAA 735
Db 765 GAGTTCAACCAAGGTTCCAGAGCTTGTGGAAGGGTGGGCGAGAACCGCTAG 824
Qy 736 CGCGTCTCGAGATGATCAACTCCTTTGATCTCTTGAAGGCACTGATCCTTGCACC 795
Db 825 CGTGACAGAGCTCCATCACTGCTTGTATCTACTTGAAGGCCCTGATCCAGCCAGC 884
Qy 796 CTTGAGAAAGTTCCTTGGAGAGATCCCATGATGTTCAATGTTGATTTCTCATCCCAAC 855
Db 885 TTGGAGAAAGTTCCTTGGACATACATGATGTTCAATGTTGATTTCTCTCCACAC 944
Qy 856 GGATATCTCGCTCAAGACATGTTTGGGGATGCCGACACCGGTGGCAGGTTGTTAC 915
Db 945 GGATATTTTGGCTCATCCATGATGTTGGGATACCTGATACCTGATGCGCAGGTTGTGAC 1004
Qy 916 ATCTTGAATCAAGTCCGAGCTTTGGAGAAATGAGATGCTCCCTCGATTAACACAGAGA 975
Db 1005 ATCTTGAATCAAGTCCGCTTGTGGAGATGAGATGCTTGTGAGGATTAACACAGAGGC 1064
Qy 976 CTCAACATCACCCCTCGAATCTCATTTATTACTAGACTTCTTCTGATGCTGCGAACA 1035
Db 1065 CTTGACATTAACCCCTTAAGATCTCATTTGTCAACAGGCTGTGGCTGATGCTGTGGAAT 1124
Qy 1036 ACATGCGGTCAAGACTTGAAGAAATATACGAAACAGAGCACTCGGATATTTCTGAGTA 1095
Db 1125 ACATGTGGCCAGCGGCTGAGAAAGTTATGTGGACTGAGCACTGACATTTCTCGTGT 1184
Qy 1096 CCGTTCAAGACAGAAAGGGAATTTGTCGAAATGAGATCTCAAGATTTGAAAAGTCTGG 1155
Db 1185 CCATTTCAGAACTGAGAGGGGATCTCCGTAAGTGAATCTCTGTTTG--ATGTCTGG 1241
Qy 1156 CCATATCTTGAAGAACTTACACAGAGATGTTGCTCATGAAATCTCCAAAGATGTCAGGC 1215
Db 1242 CCAATACCTGAGAGACTTACACGAGAGATTTGCAAAAGAACTCATGAGAGAAATGCAAGCC 1301
Qy 1216 ACGCCAGATCTGATCATGAGAAACACAGCAGCGCAATATCTGCTCTCTGCGCA 1275
Db 1302 AAGCTGATCTCATATTTGCGAACAATACAGTACGGTAACTTGTGCGCACTCTCTGCG 1361
Qy 1276 CATAAATTAGGTTCACACAGTGCACCATGCGCATGCTTTGGAGAGACAAATATATCCA 1335
Db 1362 CATATGTTGGAGATTACCCAGTGCACCATTCGCCATCTTTGGAGAAACAAAGTACCCA 1421
Qy 1336 GATTCAATATCTAATTTGAAGAAAGCTTGAAGACAAATATCAATTTCTTGGCAATTTACA 1395
Db 1422 AACTCAGACATATATTTGACAATTTGACAGGCAATATCAATTTCTCATGCTCAGTCACT 1481
Qy 1396 GCTGATCTTTTGGCAATGAACCATACAGATTTTCAATCATCACAGTACTTTCCAGGAATT 1455
Db 1482 GCTGACCTTTATGCAATGAATCAACATGATTTTCAATCAACAGCAATTCACAGAAATT 1541
Qy 1456 GCAGAGCAAGGACACTGTTGGTCAATTAAGAGCCACACTGCTTCACTCTCTGGT 1515
Db 1542 GCTGAGACAAAGACAGACCGTGGCCAGTATGAGTCTACATTTGCTTCAACCTCTCGAT 1601
Qy 1516 CTACCGGTGTTGATACATGATATGATGTTGATCCCAATTTCAATTTGTTCCCT 1575
Db 1602 CTACCGGTGTTGATACATGATATGATGTTGATCCCAATTTCAATTTGTTCCCT 1661
Qy 1576 GGTGCTGATATGAGATATATCTTCCCTTACACCGAAGAGAGCGGAGCTTGAAGATTC 1635
Db 1662 GGAGCAGATATGACTGTCTACTTCCCATACATGAACTGACAGAGGCTTACTGCTTC 1721
Qy 1636 CATCTGAGATCGAAGACCTTCTTTACACCAAGTTGAGATGAAACCTTATGTGT 1695
Db 1722 CACCTGAATTTGAAGAGCTCTTTTACAGGATGTTGAGAACTTGAACACAAATTTGTA 1781
Qy 1696 CTCAATACCGCAACAAGCAATTTCTGTTCAATGCAAGGCTTGAATGCTGTCAAGAAC 1755
Db 1782 TTGAAGACAAAGAACAGGCTATCATCTTCAATGAGCTGTGTGACCGGTGAAAGAC 1841
Qy 1756 TTAAACGAGCTCGTCAAGTGTGTGCGGCAAGAACCCAAAGTTGCGTGTGATCTC 1815
Db 1842 ATGACAGGCTTGTGAGATGTTGCTGTAAGAAATGCTCATGTAAGATTTGGCAACCTT 1901
Qy 1816 GTAGTTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1875
Db 1902 GTGATTTGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1961
Qy 1876 AAGAAATGTTTGAAGCTGATGACAAAGTACAACTTGAACGCGCAATTCAGATGATATCA 1935
Db 1962 AAAAGATGATACAGTCTCATTTGAGAGATACAGTTGAGGGCCATATCCGGTGAATCTCA 2021
Qy 1936 TCTCAATGAAACAGAAATCCGAAATGTTGAACTTTACGATATCATTTGCGACACGAAAGT 1995
Db 2022 GCTCAGATGAAACGCTGTTCCCAATGACAGAGTTGATGACGCTATCTTGTGACACCAAGGC 2081
Qy 1996 GCGTTTACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2055
Db 2082 GCATTTGTTACGCTGATTTCTATGAGACATTTGGCTGATGCTGTTGCGAGGCAATGACA 2141
Qy 2056 TGGGTTTGGCAACTTTCGCAACTGTATACGATGAGACAGCCGAGATTAATTTGTCATGG 2115
Db 2142 TGTGTTTGGCCCAATATGAGAGATGACATGATGATGATGATGATGATGATGATGATGAT 2201
Qy 2116 AAATCTGTTTCAACATTAATCTTACATGATGATGATGATGATGATGATGATGATGATGAT 2175
Db 2202 GCTCTGATTTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2261
Qy 2176 TTTCTTGAAGATGTAAGAAAGATCCATCTCACTGAGATTAAGATCTCCCAAGAGGCTTG 2235
Db 2262 TTTCTTGAAGATGTAAGAAAGATCCATCTCACTGAGATTAAGATCTCCCAAGAGGCTTG 2321
Qy 2236 AAAAGATGAGAGAAATTAATGATGAGATTAATCTGAGAGACATTAATGATGATGATGAT 2295
Db 2322 AAGAGATTTATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2381
Qy 2296 GAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2355
Db 2382 GGTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2441
Qy 2356 CTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2415
Db 2442 CTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2501
```

OY 2416 GA 2417
DB 2502 GA 2503

RESULT 11

AA66090
ID AAC66090 standard; cDNA; 2494 BP.

AC AAC66090;

DT 13-FEB-2001 (first entry)

DE Rice sucrose synthase cDNA sequence.

XX Cold resistance; transgenic plant; sucrose decomposition; rice;
KM sucrose synthase; 89.

XX Oryza sativa.

XX JF2000245279-A.

XX 12-SEP-2000.

XX 01-MAR-1999; 99JP-00052102.

XX 01-MAR-1999; 99JP-00052102.

XX (MITA) MITSUI CHEM INC.

XX MPI: 2000-675173/66.

DR P-PSDB; AAY85664.

PT Novel method for the reinforcement of cold resistance in a plant
comprising introducing a vector encoding an enzyme that decomposes
sucrose into the plant.

PS Claim 1; Page 8-11; 22pp; Japanese.

XX This invention relates to a method for the reinforcement of cold
resistance in a plant, comprising introducing an expression vector having
a DNA encoding an enzyme for decomposing sucrose connected downstream to
a promoter expressible in the plant, and expressing the enzyme in the
plant body. Included in the invention are an expression vector used in
the method, a transformed plant carrying the expression vector, and a
transformed rice carrying the above expression vector. The method is used
for reinforcing cold resistance in a plant. The present sequence
represents cDNA encoding sucrose synthase used in the method

XX Sequence 2494 BP; 622 A; 599 C; 628 G; 645 T; 0 U; 0 Other;

Query Match 49.5%; Score 1300.6; DB 3; Length 2494;

Best Local Similarity 71.9%; Pred. No. 0;

Matches 1728; Conservative 0; Mismatches 665; Indels 9; Gaps 2;

OY 16 CTCACCTCCGCTCCACAGCTCTCCGAGCGTTGATGAGCCCTTCTTCACAGGAC 75
DB CTAGCTCCCTCCACAGCTCTCCGAGCGCTCCGAGCGCTCTCTCTATCCCAAT 95
OY 76 GAGATTTGGCGCTGCTCAAGATGAGGCGAAGGAAAGAAATCTGCAACCAT 135
DB GAGTGTATGACCTTCTCTAGGTATGTTAACAGGAAAGGAAATCTCAAGCTAC 155
OY 136 CAATATATCTAGAGTTGAAGCTATCCCTGAAAGAAAGAAAGAGCTGCTAATG 195
DB CAGCTGCTGGGAGGATTCGATGCTGATGCAAGCTGCAAGAAAGAA-----ATATCT 209
OY 196 GCATTTTGAAGTATGGAAGCTAGTCAGAAAGAGATCGTTCCTTCACATGGTTCA 255
DB CCGTTTGAAGCATCTCCGGGCTGCTCAGAAAGCATGTGCTCCCGCTGGGTTCA 269
OY 256 CTGCTGTTGTCGAAGGCGTGGTGTGGAGTACATTAGAGTGAATGTTACGCGCTT 315

DB 270 CTGCGCATCAAGCGCCAGAGCGCTGGTGTCTGGGACATCAATCGGGTGAATGTAGTGG 329
OY 316 GTTGTGAGAACTCACTGTGCTGAGTATCTCCACTCAAGAAAGAGCTGTGAGAGA 375
DB GAGTGTGAGAAAGCTGAGTGTGAGTACTGGCATTCAGAAAGAGCTGTGAGAGA 389
OY 376 AGTTCAAATGGAACCTTGTGTTTGAATGGAATTTGAGCCCTCACTCAATCCCTCC 435
DB CACACCAAGCAACTTGTGTTTCTGAGCTGATTTTGAAGCCCTCAATGCTCTCCCG 449
OY 436 CGCCCACTCTTCAAAATCCATTTGTAATGTGAGATTCCTAAATCGTCACTTTGG 495
DB CGCCGTCATGTCAGATGCTGAGAAATGGGGTGCAGTTCCTAACCGCTCATCTGCG 509
OY 496 GCAAAATTTGTCATGACAGAGAGATGACCCCTTCTCGAAATTCCTCAAGTCCAT 555
DB TCCAAATTTGTCAG 569
OY 556 TGTCAAGAGGCAAGAAATGATGTTGAATGACAGAAATTCAGAACTTGAATGCTCTCA 615
DB AACCAAGAGGCAAGAAATGATGTTGAATGACAGAAATTCAGAACTTGAATGCTCTCA 629
OY 616 CATGTTTGAAGAAAGAGAGAGATCTTGTGTAACCTTCTCTGAGACATGTCG 675
DB TCATCCCTTGAAGAAAGAGAGAAATTCATGAGGCAATCTTCAAGACAGCCCTTCTG 689
OY 676 GAATTCGAACACCGGTTCCAGAAATGCTTGAAGAAAGAGAGAGAGAGAGAGAGAG 735
DB GAGTTCAACCAAGAGTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 749
OY 736 CGCGTGTGAGAGATGATCAATCTCTTGTGATCTTCTGAGAGCAATGATCTTTCAC 795
DB CGTGTGCTTGAACATATCACTTCTCTTGAATCTTCTTGAAGCCCTTGAAGCCGAC 809
OY 796 CTTGAGAGATTCCTTGGAGAGATCCCAATGCTTCAATGTTGATTCATCTCCAC 855
DB TTGAGAAAGTCTTGGAACTATTCATATGTTGATGTTATCTCTCTCCGAT 869
OY 856 GGAATCTTCCGCTCAAGACATGTTTGGGGTATCCCAACCGGTTGGGAGGTTTAC 915
DB GGAATCTTCCGCTCAAGACATGTTTGGGGTATCCCAACCGGTTGGGAGGTTTAC 929
OY 916 ATCTTGATCAAGTCCGAGCTTGGAGATGAGATGCTCCCTCGATTAACCAAGAGA 975
DB ATTTTGAACCAAGTCCGAGCTTGGAGATGAGATGCTCCCTCGATTAACCAAGAGA 989
OY 976 CTCAACATCAACCCCTCGATCTCTATTAATGATCTTCTCTGATGCTGGAACA 1035
DB CTGATATCAACCTTAAGATCTCTATGTAACAGGGCTGTTGCTGATGCTGTTACT 1049
OY 1036 ACATCCGCTCAAGATCTTGAAGAAATGATACGAAACAGAGCACTGGATTTCTCGAGTA 1095
DB ACATCCGCTCAAGATCTTGAAGAAATGATACGAAACAGAGCACTGGATTTCTCGAGTA 1109
OY 1096 CCTTCGAAGAAAGGAAATGTTGGAATGAGATGCTCCCTCGATTAACCAAGAGA 1155
DB CCTTCGAAGAAATGTTGGAATGAGATGCTCCCTCGATTAACCAAGAGA 1166
OY 1156 CCATATCTGAAACCTTCAACAGAGAGATGCTCATGAAATCTCCAAAGAGTCAACGCG 1215
DB CCATATCTGAAACCTTCAACAGAGAGATGCTCATGAAATCTCCAAAGAGTCAACGCG 1226
OY 1216 AGCGAGATCTGATCATCGGAAACACAGAGAGGCAATATGTCGCTCTGCTGCA 1275
DB AAACCTGATCTCATCATATGAGCAATTAACAGTATGAAACCTTGTGCACTCTGCTGCT 1286
OY 1276 CATAAATTAAGTGTCAACAGTGCACCATGCGCATGTTGAGAGAGCAAAATATATCA 1335
DB CACAAATTAAGTGTCAACAGTGCACCATGCGCATGTTGAGAGAGCAAAATATATCA 1346
OY 1336 GATTCAATATCTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1395

Db 1347 AACTCAGCATATATCTGACAGAGTTTGACACCCAGTACCACTTCTCATGCCAAATTCAC 1406
 QY 1396 GCTGATCTTTTGGCATATGACCATATCAGATTTTCATCACCAGTACTTTCCAGGAAT 1455
 Db 1407 GCTGATCTTATCGGCATATACACCTGATTTTCATCATCAGTACATTCAGGAAT 1466
 QY 1456 GCAGAGCAAGAGACACTGTTGGTCAATACAGAGCCCACTGCTTTCATCTTCTGCT 1515
 Db 1467 GCTGAGAGCAAGACACTGTTGGGCGATATATCACAATGATTCACCTTCTGCTG 1526
 QY 1516 CTCTACCGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1575
 Db 1527 CTCTACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1586
 QY 1576 GGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1635
 Db 1587 GGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1646
 QY 1636 CATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1695
 Db 1647 CACCTGAAATTTGAGAGCTTCTCTACAGTGAAGTCGAAACGATGAAACAAAGTTGTA 1706
 QY 1696 CTCAATGACCCGCAACAGCCAAATCTGTTCACAAAGCCAGCTGATGATGATGATGAT 1755
 Db 1707 TTGAAGGCAAGAACAGCCCAATCTTCTCAGTGCCTGCTGATGATGATGATGATGAT 1766
 QY 1756 TTAACCGGACTCTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1815
 Db 1767 ATGACAGGCTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1826
 QY 1816 GTAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1875
 Db 1827 GTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1886
 QY 1876 AAGAAAAGTTTGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1935
 Db 1887 AAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1946
 QY 1936 TCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1995
 Db 1947 GCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2006
 QY 1996 GCTTTTGTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2055
 Db 2007 GTCTTTGTCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2066
 QY 2056 TGGGTTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2115
 Db 2067 TGTGTTTCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2126
 QY 2116 AATCTGTTTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2175
 Db 2127 GTGTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2186
 QY 2176 TTTCTTTGAAAGTGTAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 2235
 Db 2187 TTTCTTTGAAAGTGTAGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 2246
 QY 2236 AAAAGAAATGAGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2295
 Db 2247 CAGAGGATTTTACGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2306
 QY 2296 GGAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2355
 Db 2307 GGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2366
 QY 2356 CTGAGATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2415
 Db 2367 ATTGAGATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2426
 QY 2416 GA 2417
 Db 2427 GA 2428

RESULT 12
 ID ADC68523
 AD 68523 standard; cDNA; 2825 BP.
 XX
 AC ADC68523;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE S. arundinaceus fructan biosynthesis protein cDNA SEQ ID NO:15.
 XX
 KM Lolium perenne; Festuca arundinacea; lignin; fructan; tannin;
 KM biosynthetic pathway; plant; gene; ss.
 OS Schedonorus arundinaceus.
 XX
 FN W02003040306-A2.
 XX
 PD 15-MAY-2003.
 XX
 PF 07-NOV-2002; 2002WO-NZ000239.
 XX
 PR 07-NOV-2001; 2001US-0337703P.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 PA (WRIG-) WRIGHTSON SEEDS LTD.
 PI Demmer J, Forster RL, Gibson JB, Shenk MA, Norris MG, Glenn M;
 PI Saulsbury KM, Hall C;
 XX
 DR WPI; 2003-441544/41.
 DR P-PSDB; ADC68367.
 XX
 PT New polynucleotide encoding polypeptides from Lolium perenne or Festuca
 PT arundinacea, useful for modulating the biosynthesis of lignin, fructan or
 PT tannin in a plant.
 PS
 PS Claim 1; SEQ ID NO 15; 240bp; English.
 XX
 CC The present invention describes isolated polynucleotides (i) encoding
 CC proteins (ii) from Lolium perenne and Festuca arundinacea which are
 CC active in lignin, fructan and tannin biosynthetic pathways. Also
 CC described: (1) an isolated oligonucleotide probe or primer comprising at
 CC least 10 contiguous residues complementary to 10 contiguous residues of
 CC (1); (2) a kit comprising the oligonucleotide probe or primer; (3) a
 CC genetic construct comprising (1); (4) a transgenic plant cell comprising
 CC the genetic construct of (3); (5) a plant or its seed, fruit or progeny
 CC comprising the transgenic plant cell of (4); (6) modulating one or more
 CC of the lignin, fructan or tannin compositions of a plant; (7) producing a
 CC plant having one or more of the lignin, fructan or tannin compositions;
 CC and (8) modifying the activity of (ii) involved in a lignin, fructan or
 CC tannin biosynthetic pathway in a plant. (i) can be used for modulating
 CC the biosynthesis of lignin, fructan or tannin in a plant. The present
 CC sequence is used in the exemplification of the present invention.
 XX
 SQ Sequence 2825 BP; 678 A; 721 C; 726 G; 700 T; 0 U; 0 Other;
 Query Match 49.4%; Score 1296.6; DB 9; Length 2825;
 Best Local Similarity 71.7%; Pred. No. 0;
 Matches 1727; Conservative 0; Mismatches 675; Indels 6; Gaps 2;
 QY 7 GAGGCTCTCTCACTCGCTGTCACAGTCTCGTGAAGCGTTGGATGAGACCTTCTTCTGCT 66
 Db 121 GAGGCTCTCTGAGCCCTCTCCACAGGTAAAGGAGGTATCGCGCATTCCTCTCAGCG 180
 QY 67 CACAGCAAGATTTTGGCTTGTCTCAAGATGAGGCGCAAGGAAAGGAATTTCTG 126
 Db 181 CACACCAATAGAGTCGCGCCGCTTCTCAAGGCTTTTAACATGAGAAAGGGAATCTT 240
 QY 127 CAACACATCAATTAATTTAGAGT---TTGAGCTTCTCCTGAAGGACAGAAAGAG 183
 Db 241 CAGCCCAAGATCACTGCTGATGATCAATGCTGATTTCCGAGGCTGAGCGAGAGAG 300

```
OY 184 CTCGCTATGAGCATTTTGAAGTATTTGAAGGCTAGTCAGAAAGCATGCTGTTGCT 243
DB 301 CTGAAGATACCGCTTTGAGATCTCTMAAGGGGGGACAGAGAGCAATTGTATCCCT 360
OY 244 CCATGGGTGACCTGCTGCTGCTCAAGGCGCTGCTGTTGGAGTACATTAGAGTAT 303
DB 361 CCATGGGTGCTCTGCGCATCCGGCAAGGCTGGCGTCTGGAGTATGTGAGGGCTCAT 420
OY 304 GTTACGCGCTTGTGTGAGAACTCAGTGTGTGATCTCCACTTTCAGGAAGAG 363
DB 421 GTGAGCGAGCTCGGGGTGAGAGTGTAGTCTCTGAGTATTTGCAATTCAMGAGACA 480
OY 364 CTGTGTGAGGAAGTTCAATGAAATCTTTGTTGAAATGGAATTTGAGCCCTTCAAC 423
DB 481 TTGGTGAATGAAGACATGATTAACAATTGTGCTTGAAGCTGAGCTTTGGGCAATTCAAT 540
OY 424 TCATCATTTCCCGCCCAACTCTTCAAAATCCATTTGTAATGTGTGAGAGTTCTTAAT 483
DB 541 GCGTCTTCCCGCCGCAATCGCTGTGCAAGTCTATTTGCAATGTGTGCAAGTCTTGAAC 600
OY 484 GGTCACTTTGGGCAAAATTTGTCATGACAGAGAGACATGCACTTTGTGCTGAATTC 543
DB 601 AGGCACCTGTATCAAACTGTTCCATGACAGAGAGACATGACCTTGTCTCAACTTC 660
OY 544 CTCAGAGTCAATGTACAGAGGGCAAGAACATGATTTGAATGACAGAAATCAGAACTTG 603
DB 661 CTCCTGTGACACAACTTACAGGGGATGACATGATGTGTGACAGAGAAATTCGACGTTC 720
OY 604 AATGCTCTTCAACATGTTTGAAGAAAGAGAGATCTTGTGATCCCTACCTCTGAG 663
DB 721 AGTGCCCTTCAAGATGACATGAGAGAGAGAGAGATCTGTGACGGCTTCCAGCTAC 780
OY 664 ACACCATGTCCGAATTTGAACAACCGGTTCCAGAAATCGGTTTGAAGAGGTTGGGT 723
DB 781 ACCCGTACTCAGAGTTCCACACCGGTTCCAGAACTGTGTTGAGAAAGGTTGGGGC 840
OY 724 GACACCGAGAACGGGTGCTGAGATGATCCAACTCTTTTGGATCTTCTTGAAGCACT 783
DB 841 GACTGTACTACGCGTACACAGAGACTATTCACCTTCTTGTGACCTTCTCGAGGCCCT 900
OY 784 GATCCTTGACCCCTTGAGAGATCTTGTGGAGAAATCCCATGATGTTCAATGTTGTGAT 843
DB 901 GATCATATCACTTGTGAAAGTTTCTCGGAACAATCCCATGATGTTCAATGTTGTGAT 960
OY 844 CTCACTCCCAACGATCTGCTGCTCAAGCAATGTTTGGGATTCGCGACACCGGTGG 903
DB 961 CTCTCTCTCATGTGTACTTGTGCTCAGGCAATGTCTTGGGGTACCCAGACACTGTGTGG 1020
OY 904 CAGGTTGTTTCAATTTGAGATCAAGTCCGAGCTTGGAGATGAGATGCTCTCCGTATTA 963
DB 1021 CAGGTTGTTTCAATTTGAGATCAAGTCCGAGCTTGGAGATGAGATGCTCTATTTGAATTC 1080
OY 964 AAGCAACAAGACATCAACATCAACCCCTGGAATCTCATTTACTAGACTTCTTCTGAT 1023
DB 1081 AAGCAACAAGACATCTTGAATTAACCAAAAGATTTAATAGTACACAGGTGTCTCTGAT 1140
OY 1024 GCTGTGGAACAACATGCGGTCAACGACTTGAAGAAAGTATCGAAACAGAGCATCGGAT 1083
DB 1141 GCACACGGCACCACTGCGGCGCAGCGCTTGAAGAGTCTTGGACATGAGACACCCAC 1200
OY 1084 AATTCCTGAGATCCCTTCAAGAACAGAAAGGAAATTTGCAAAATGGAATCTCAAGATTT 1143
DB 1201 AATTCGCGGTGCAATTCAAAACAGAAAGTGGATTTGTGCAAAATGGAATCTCAAGATTT 1260
OY 1144 GAAAAAGTCTGGCCATTTACTGGAACCTTACACAGAGAGATGTTGCTCATGAATCTCCAA 1203
DB 1261 G---AAGTCTGGCTTTACTGGAACCTTACACAGATGATGTGGACATGAGATTTCTGGA 1317
OY 1204 GAGTTGACAGGACCGCAGATTTGATCATCGAAACACAGGACGCAATATCGTGGC 1263
DB 1318 GAGCTGACAGGCAACCTGATCTTGTGATCATTTGAAACTACAGATGAGAAACCTTGTGGC 1377
OY 1264 TCTTGTCTCGACATAAATTAGTGTGACACAGTGCACCATGCGCCATGCTTTGAGAG 1323
DB 1378 TGTTTGTGGACACAAGTTGGGGTTACTCATTTGACATTTGCTCATGCGCTTAGAGAA 1437
OY 1324 ACMAAATATCCAGATTCAGATATCTATTGGAAGAGCTTGAAGCAAAATACATTTCT 1383
DB 1438 ACMAAGTACCCCACTGACCTTACTGGAAGAAATTTGAGATCACTACACCTTTCC 1497
OY 1384 TGCAATTTACAGGTGATCTTTTGAATGAAACATACAGATTTCAATCATCACCACT 1443
DB 1498 TGCAATTTACAGGTGATCTTTTGAATGAAACATGATGATCTTCAATCATCACCACT 1557
OY 1444 TTCCAGAAATTTGAGAGAGAGAGACACTGTTGTCATATCAGAGGCACTGCTTC 1503
DB 1558 TTCCAGAGATTTGCTGGAACAAGAGATCTGTAGGGAGATATGATGTCGACATGGAATTC 1617
OY 1504 ACTTCTCTGTCTCTACCGTGTGTGATGATGATATGATGATGATGATGATGATGATGAT 1563
DB 1618 ACAATGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1677
OY 1564 ATTTGTTCCCTGCTGCTGATATGAGATATATCTTCCCTTACACGAGAGAGCGGAG 1623
DB 1678 ATGCTCTCTGCTGCTGATATGAGATATATCTTCCCTTACACGAGAGAGCGGAG 1737
OY 1624 TTGAAGCATTTTCACTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1683
DB 1738 CTTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1797
OY 1684 CACTTATGTGTCTCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1743
DB 1798 CACAAATTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1857
OY 1744 GGTGCAAGAGCTTACCGGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1803
DB 1858 GGTGCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1917
OY 1804 TTGCTAATCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1863
DB 1918 CTGATTAATCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1977
OY 1864 AAGGCTGAATGAAGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1923
DB 1978 CAGGCGAGGTTCAAGAGAGATTTGATCTTATCAGAGAGAGAGAGAGAGAGAGAGAG 2037
OY 1924 AGATGATATCATCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1983
DB 2038 CGTGAGATCTGCTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2097
OY 1984 GACAGAAAGGCTTGTGACAGCTGATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 2043
DB 2098 GACATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2157
OY 2044 GAGGCAATGACTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2103
DB 2158 GAGGCAATGACTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2217
OY 2104 ATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2163
DB 2218 ATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2277
OY 2164 ATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2223
DB 2278 CTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2337
OY 2224 CAAGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2283
DB 2338 CAGGCGAGATCTCAGGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2397
OY 2284 TTGACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2343
DB 2398 ATGACCTCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2457
OY 2344 AGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2403
```


|||||
Db 2458 ACTGCTCCTACCTTGAGATGCTTACGCTCTCAAGTACCGCAAAATGCTACCAACCGTT 2517
QY 2404 CCATTGGC 2411
Db 2518 CCATTGGC 2525
RESULT 13
AD68521
ID AD68521 standard; cDNA; 2754 BP.
XX
AC AD68521;
XX 18-DEC-2003 (first entry)
DE S. arundinaceus fructan biosynthesis protein cDNA SEQ ID NO:13.
XX
XX Lolium perenne; Festuca arundinacea; lignin; fructan; tannin;
KM biosynthetic pathway; plant; gene; ss.
XX
XX Schedonorus arundinaceus.
XX
XX WO2003040306-A2.
XX
XX 15-MAY-2003.
XX 07-NOV-2002; 2002WO-NZ000239.
XX 07-NOV-2001; 2001US-0337703P.
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (WRIG-) WRIGHTSON SEEDS LTD.
PI Demmer J, Foster RL, Gibson JB, Shenk MA, Norriss MG, Glenn M;
PI Saulsbury KM, Hall C;
XX WPI: 2003-441544/41.
DR P-PSDB; AD68365.
XX
PT New polynucleotide encoding polypeptides from Lolium perenne or Festuca
PT arundinacea, useful for modulating the biosynthesis of lignin, fructan or
PT tannin in a plant.
XX
XX Claim 1; SEQ ID NO 13; 240bp; English.
XX
XX The present invention describes isolated polynucleotides (I) encoding
XX proteins (II) from Lolium perenne and Festuca arundinacea which are
XX active in lignin, fructan and tannin biosynthetic pathways. Also
XX described: (1) an isolated oligonucleotide probe or primer comprising at
XX least 10 contiguous residues complementary to 10 contiguous residues of
XX (I); (2) a kit comprising the oligonucleotide probe or primer; (3) a
XX genetic construct comprising (I); (4) a transgenic plant cell comprising
XX the genetic construct of (3); (5) a plant or its seed, fruit or progeny
XX comprising the transgenic plant cell of (4); (6) modulating one or more
XX of the lignin, fructan or tannin compositions of a plant; (7) producing a
XX plant having one or more of the lignin, fructan or tannin compositions;
XX and (8) modifying the activity of (II) involved in a lignin, fructan or
XX tannin biosynthetic pathway in a plant. (I) can be used for modulating
XX the biosynthesis of lignin, fructan or tannin in a plant. The present
XX Sequence 2754 BP; 652 A; 685 C; 684 G; 733 T; 0 U; 0 Other;
SQ
Query Match 49.3%; Score 1295.4; DB 9; Length 2754;
Best Local Similarity 70.2%; Pred. No. 0; Mismatches 742; Indels 9; Gaps 2;
Matches 1769; Conservative 0;
QY 16 CTCACCTGCGTCAACAGTCTCCGTGAGCGTTGATGAGACCTCTTGTCTCAGAGAAC 75
Db 125 CTGACTGCTCTCAACAGTCTCCGGAGCGCTTGGTGCACCTTCTCTCCATCCCAAC 184
QY 76 GAGATTTGGCCTTGCTCTCAAGATCGAGGGCAAGAAAGAAATTTCTGCAACACCAT 135

|||||
Db 185 GAGCTCAATGCGCTCTTTTCCAAAGTATGTTCCACAGGCAAAAGAAAGCTTCAAGCGCAC 244
QY 136 CAAATTATTTAGAGTTGAAAGCTATCCCTGAAGAGAAACAGAAAGAGCTCGTAAATGTT 195
Db 245 CAGCTGCTCACTAGAGTTTGAAGCATTTGTTGAGTCTGACAAAGAGAGG-----TATGCA 298
QY 196 GCATTTTGTGAAGTATGAAAGGCTATGACAGAGCCGATCGTGTTCCTCCATGGGTTTGA 255
Db 299 CTTTTCAGACATCTCCGCTGCTGCCAGAGCAATTTGGCTTCCCTCCATGGGTTTGA 358
QY 256 CTGCTGTGTGCTCCAGGCGGTGTTGGAGTACATTGAGTAATGTTCAAGCCCTT 315
Db 359 CTTGCCATCAAGCCAGGCGCTGCTGTCTGGACTCAATTAAGGTAATGTTAGAGAGTTG 418
QY 316 GTTGTGAGGAACTCACTGTGCTGAGTATCTCCATCTCAAGAGAGCTGTGATGGA 375
Db 419 GCGTTGAGAGAGCTGACAGTTTCCGAGTACTAGCATTTCAAGAGAGAGCTTGTGATGAG 478
QY 376 AGTTCAAAATGAAACCTTGTGTTTGAATTTGAGATTTGAGCCCTTCAACTCATATTCGCC 435
Db 479 CAGCCAGACAGCAAGTTGTGCTTGAAGCTTGAATTTGAGCCTTCAATGCTCCTCCCA 538
QY 436 CGCCCACTCTTCAAAATCCATTTGTAATGCTGTGAGTTCCTAATTCGCACTTTCG 495
Db 539 CGTCTTCATGTCAAAGTCATTCATTTGAATGAGGCTGACGTTCTTAAACCGTACCTGTCT 598
QY 496 GCAAATTTGTTCCATGACAGAGAGAGCATCCCTTGTCTGAAATTCCTCAGAGTCAT 555
Db 599 TCCAGAGCTTTCAGAGCAAGAGAGAGCTCTACCCACTACTGAACTTCTGTAAGGCCAT 658
QY 556 TGTCACAAGGCAAGAAATGATGTTGAATGACAGAAATTCAGAACTTGAATGCTCTTCAA 615
Db 659 AACCAAGAGGCAAGCAATGATGCTTAAAGACAGAAATTCAGAGCTTTCGAGGCTCCAA 718
QY 616 CAGTTTGTGAGAAAGAGAGAGTATCTGTACCTTACTCCTGAGACACATGAGCC 675
Db 719 TCAGCCCTTGAAGAGAGAGATATCTTAACAAGCATCTTGAAGACCCCGCTCTTT 778
QY 676 GAATTCGAAACCGGTTCCAGAGAAATCGTTTGAAGAGAGTTGGGTTGACACCGCAGAA 735
Db 779 GAGTTCAACACAGAGTTTCAAGAGCTTGGCTTGAAGAGAGTTGGGTTGACACCGCAGAA 838
QY 736 CGGCTGCTCGAGATGATCCAACTCTTTTGAATTTCTTGAAGCACTGATCCTTGCAAC 795
Db 839 CGTGTACAGACACCAATCTGCTTGTGATCTGTTAGAGCCCTCGATCCGACAGC 898
QY 796 CTGAGAAAGTTCTTGGGAGAAATCCCATGATGTTCAATGATTCACCTCCAC 855
Db 899 TTGAGAAAGTTCTTGAACCTATACGATATGTTCAATGTCGTTATCTGTCACAC 958
QY 856 GGAATCTTCCCTCAAGCAATGTTTGGGGTATCCGACACCGGTGGCAGGTTGTTTAC 915
Db 959 GGAATTTTGTCTCAATGATGATGTTGGATACCTGTATACCGGTGGCAGGTTGTATAC 1018
QY 916 ATCTTGATCAAGTCCGAGCTTTGAGAGATGAGATGCTCTCGTATTAAGCAAGAA 975
Db 1019 ATCTTGATCAAGTCCGAGCTTTGAGAGATGAGATGCTCTCGTATTAAGCAAGAGGC 1078
QY 976 CTCAACATCAACCCCTGAATCCATATTAAGATTTCTTCCGTAGTGCCTGAGACA 1035
Db 1079 CTTGACATTAACCCCTGAATCCATATTAAGATTTCTTCCGTAGTGCCTGAGACA 1138
QY 1036 ACATGCGGTCAACGACTTGAAGAAATGTAAGCAACAGACACTCGATATTTCTGAGTA 1095
Db 1139 ACATGTCGCAAGCGGCTGAGAGAGTTATTTGGAAGTACAGACACTGACATTTCTCGGTT 1198
QY 1096 CCTTTCAGAAACAGAAAGGAATTTGTAAGATGATTCGAAGATTTGAAAAAGTTGG 1155
Db 1199 CCAATTCAGAACTGAAAAAGGAGATCTCCGTAAAGTATCTCGTTTGG--ATGTCGG 1255
QY 1156 CCAATCTTGAAGACCTTACAGAGAGATGTTGCTCATTAATTTCTCAAGAGTTGCAAGGC 1215

Db 1256 CCATACCTGAGACATACACGAGAGTGTGCAACGAACCTCATGAGAAATGAGAC 1315
 Qy 1216 AGCCGAGATGATGATGAGAAACGACGCGCATATGCTGCTGCTGCGCA 1275
 Db 1316 AAGCTGATCTCATATTATGGAACCTACAGTACGCGTAACCTTGTGCGCACTGCTGCG 1375
 Qy 1276 CATAAATAGGATGTCACAGTGCACATCGCCCATGCTTGGAGAAACAAATATCA 1335
 Db 1376 CATAGTTGGAGTTACCAAGTGCATTTGCCCATGCTTGGAGAAACAAATATCC 1435
 Qy 1336 GATTGATATCTATTTGGAAGAGCTTGAAGCAAAATCAATTTCTCTCCCAATTA 1395
 Db 1436 AACTGACATATATTTGGAACAAATTCGACAGCCAAATCAATTTCTCATGCGAGTCA 1495
 Qy 1396 GCTGATCTTTTGGCAATGACATACATTTTCAATCACCAGTACCTTCCAGGAAT 1455
 Db 1496 GCTGACCTTATTTGCAATGATCACTGATTTGATCACCAGACATTCAGGAAT 1555
 Qy 1456 GAGGAGCAAGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1515
 Db 1556 GCTGAGACAGAGACAGGCTGAGGACGATGATGATGATGATGATGATGATGATGAT 1615
 Qy 1516 CTGATCCGTTGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1575
 Db 1616 CTGATCCGTTGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1675
 Qy 1576 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1635
 Db 1676 GAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1735
 Qy 1636 CATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1695
 Db 1736 CACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1795
 Qy 1696 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1755
 Db 1796 TTGAAGACAGAGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1855
 Qy 1756 TTAACCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1815
 Db 1856 ATGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1915
 Qy 1816 GTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1875
 Db 1916 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1975
 Qy 1876 AAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1935
 Db 1976 AAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2035
 Qy 1936 TCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1995
 Db 2036 GCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2095
 Qy 1996 GCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2055
 Db 2096 GATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2155
 Qy 2056 TGCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2115
 Db 2156 TGTGTTTGGCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2215
 Qy 2116 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2175
 Db 2216 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2275
 Qy 2176 TTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2235
 Db 2276 TTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2335
 Qy 2236 AAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2295
 Db 2336 AAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2395

Qy 2296 GAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2355
 Db 2396 GGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2455
 Qy 2356 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2415
 Db 2456 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2515
 Qy 2416 GAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2475
 Db 2516 GACGCGAGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2575
 Qy 2476 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2535
 Db 2576 GTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2635
 RESULT 14
 ADA71337
 ID ADA71337 standard; DNA; 2450 BP.
 AC ADA71337;
 DT 20-NOV-2003 (first entry)
 DE Rice gene, SEQ ID 4660.
 KW Plant; bacterial infection; fungal infection; viral infection; rice;
 gene; ds.
 OS Oryza sativa.
 PN WO200300898-A1.
 PD 03-JAN-2003.
 PF 22-JUN-2001; 2001WO-1B001105.
 PR 22-JUN-2001; 2001WO-1B001105.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katsagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 DR WPI; 2003-175290/17.
 PT Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 PS Claim 6; SEQ ID NO 4660; 899bp; English.
 CC The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 SQ Sequence 2450 BP; 556 A; 674 C; 628 G; 592 T; 0 U; 0 Other;
 Query Match 49.1%; Score 1287.6; DB 7; Length 2450;
 Best Local Similarity 71.4%; Pred. No. 0;
 Matches 1724; Conservative 0; Mismatches 665; Indels 6; Gaps 2;
 Qy 7 GAGGCTCTCATCGGCTCCAGATCTCGTGAAGCTTGTGATGAGACCTTCTGCT 66

```
Db 19 GACGCGCTGAGCGCCCTCCACAGGCTCAGGAGCGCATCGGCACTCCCTCTCGCC 78
Qy 67 CACAGGAACGAGATTTGGCTTGGCTCTCAAGATCGAGGGCAAGAAAAGAAATTCG 126
Db 79 CACCCCAATGAGCTGTGCGCGCTCTTCAACAGGCTGGTTAACCTCGGAAAGGAATGTT 138
Qy 127 CAACACATCAATATTCTAGAGT--TTGAAGCTATCCTGAAGAACAGAAAGAG 183
Db 139 CAGGCGCCACAGATCATTTGAGTACAAACGCAATCTCGAGGCTGATCGTAGAAG 198
Qy 184 CTCGTATAGTGAATTTTGAAGTATTAAGGCTATCAGGAAGGATGTGTGCT 243
Db 199 CTGAAGAGCGGCTTTTGAAGATGCTGAAGAGTSCAAGAGAGGATTTGTTATCTCC 258
Qy 244 CCATGGTGTGACATTTGCTGTTCTCAAGGCTGGTGTGGAGTACATTAGATGAAT 303
Db 259 CCAATGGTGTGCTTTGCCATCGCGCCGAGGCCGGTGTCTGGAGTATGTAAGGTCAAT 318
Qy 304 GTTACGCGCTTGTGTTTGAAGAACTACTGTGTGATCTCCACTTCAAGAAAG 363
Db 319 GTGAGCGAGCTCGCTGTGATGTGCTGACAGTCCCGAGTACTTGCAGTTCAAGAA 378
Qy 364 CTGTGTATGTAAGTTCAAATGGAACCTTTGTGGAATTTGAAATTTGACCCCTTCAAC 423
Db 379 CTGTGGAAGAGGACCAACAACTTTGTGCTTGAAGCTGGAATTCGAGCCATTCAT 438
Qy 424 TCATCATTTCCCGCCCAACTTTTCAAAATCATTTGTAATGGTGTGAAATTCCTAAT 483
Db 439 GCGTCTTCCCTCGTCTTCTGTCTGTCAAGTCAATTTGGCAATGGTGTGCAATTCCTCAAC 498
Qy 484 CGTCACTTTTGCGCAAAATTTGTTCTATGACAGAGAGCATGACCTTTGCTGGAATTC 543
Db 499 AGGATCTGTATCGAACTCTTCATGACAAAGGAGACATGATCCCACTCTCAACTTC 558
Qy 544 CTCGAGTCTATGTGCAAGGGCAAGAACATGATGTGAATGACAGAAATTCAGAACTTG 603
Db 559 CTGTGACACAACTACAAAGGGCATGACATGATGTGAATGACAGAAATTCGAGTCTC 618
Qy 604 AATGCTCTTCAACATGTTTGAAGAAACAGAGATCTTGGTACCTTACTCTCTGAG 663
Db 619 AGTGTCTCCAAAGGTCTCTGAGAGGCTGAGAGCATCTTTCGTCTTTCAGCAGAC 678
Qy 664 ACACCATGTGCGGAATTTGAAACACCGGTTCCAGAAATCGGTTTGAAGAGTTGGGT 723
Db 679 ACCCTTACTCAGAAATTCACCAACAGGTTCCAGGAATTTGCTGGAAGAGGTTGGGT 738
Qy 724 GACACCGGAAACGCGTCTGAGATGATCCAACTCTTTTGGATCTTTGAGCAACT 783
Db 739 GACTGTGCAAGCGTTCCAGAGACCAATTCACCTCTTTTGAACCTTCTTGAAGGCCCT 798
Qy 784 GATCTTTCGACCCCTGAGAAATCTTGGAGAAATCCCATGATGTTCAATGTTGTGAT 843
Db 799 GATCGTTCACCTCTGAGAAATGTTCTTGGAAACATCCCAATGGTGTTCATGTTGTCAAT 858
Qy 844 CTCACCTCCCAACGATACTTGGCTCAGAACAAATGTTTGGGATATCCGACACCGGTGG 903
Db 859 ATGTCTCCCAACGATTTACTTGGCGCAACCAATGTCTTGGGTTACCTCGAACATCGAGGG 918
Qy 904 CAGGTTGTTTCACTTTGGAATCAAGTCGGAAGCTTTGGAATGAGAGTCTCTCCGATTA 963
Db 919 CAGGTTGTCTCACTTTTGAACCAAGTCGCTGTATGGAATGAGAGTCTTGAAGATC 978
Qy 964 AAGCAACAGAGACTCAACATCAACCCCTGAAATCCATTTACTATGACTTCTTCTGAT 1023
Db 979 AAGCAGCAGAGTCTCAATTTTACACAGCATCTTATTTGTCACAGGTTCTCTCCGAT 1038
Qy 1024 GCTGTGGAACAACATGCGGTCAACGACTTGAAGAAAGTATACGGAACAGACACTCGAT 1083
Db 1039 GCAACTGGGCAACACTGCGGTCAAGGCTTGAAGAGCTTGGCACCGAACAACCCAC 1098
Qy 1084 ATTCTTGAGTACCTTTCAGAACGAAGGAATTTGTTGAAATGGAATTCGAAGATT 1143
Db 1099 ATCTTACGTGTGCTTTCAAGAACTGAAATGGAATTTGTGCAAGTGAATCTCAGTTT 1158
Qy 1144 GAAAAGTCTGGCCATATCTTGAAGAACTTACACAGAGAGTGTGCTCATGAATCTCCAA 1203
Db 1159 G---AAGTCTGGCGGTACTGTGAAGAACTTCACTGATGATGTGGGCAAGAAATGTGGA 1215
Qy 1204 GAGTTGACGGCACGCCAGATCTGATATCGGAACAAACAGCGACGCAATATGTGGCC 1263
Db 1216 GAGTCTCAGGCAATCTTGATCTGATATCGGAACATCAAGTACGGAACCTTGTGTGA 1275
Qy 1264 TCTTGTCTGCAATAAATTAAGGTGACACAGTGCACATCGGCCATGCTTGGAGAG 1323
Db 1276 TGCTTGTCTGCAACAGATGGGTGTTACCAATTTGACATTTGCCATGCGCTTGAGAA 1335
Qy 1324 ACAAATATTCAGATTCATATCTATTTGAAGAGCTTGAAGACAAATACATTTCTCT 1383
Db 1336 ACCAATATCCCAACTGTGACCTTACTTGAAGAAAGTTTGAAGATCACTACATTCCTCA 1395
Qy 1384 TGCCATTTTACAGTGAATCTTTTGCATGACATACATATTTTCAATCATCACTAAT 1443
Db 1396 TGCCAGTTCAACTGACTTGAATTCGATGAACACGCTGACTTCAATCATCACTAATCC 1455
Qy 1444 TTCCAGGAATTCGAGAGCAAGAACACTGTTGTGCAATTCAGAGCACTACTGTTTC 1503
Db 1456 TTCCAGAGATTCGCGAACAAGAGACACTGTTGTGATGATGATCTCATATGGCAATTC 1515
Qy 1504 ACTTTCCTGCTCTACACGCTGTGTGATCATGTATCATGATGTGTTGATCCCAATTCAC 1563
Db 1516 ACCATGCTGCTGTGTGATCCGATGTTGTCACGATATGATGTTTGAACCGGAATTTAAC 1575
Qy 1564 ATTGTTCCTGCTGTGTGATATGAGATATATCTTCCCTTACACCGAAGAGAGGAGG 1623
Db 1576 ATTGTCTGCTGTGTGAGCATGTCCATCTACTTCTTACTGTGATCTGTGATAGG 1635
Qy 1624 TTGAAGATTTTCACTCTGAGATGAAAGACCTTCTTACACAAAGTTGAATGAAGAA 1683
Db 1636 CTCACCTCCCTCCACCAAGATTTGAGAGTGTCTTACAGCAAGTTGACACAAAGAG 1695
Qy 1684 CACTTATGTGTCTCAATGACCGCAACAGCAATTCCTTGTTCACATGCAATGCCAATTCAT 1743
Db 1696 CACAAATTTATGTGAGAGACAGAAACAGCAATCATCTTCTCATGAGGCTGTGCTTGAAC 1755
Qy 1744 CGTGTCAAGAACTTAAACCGGATCTGTGAGTGTGCGGCAAGAACCCAAAGTTCGTTGAG 1803
Db 1756 CGTGTCAAGAACTTGAATGTGTGTGAGTGTGAGTGTGCAACCTCGCGTGCAGAGAG 1815
Qy 1804 TTGGCTAACCTCGATGTGATGAGTGTGATAGGGAAGAAATCTTAAGATTTGAGAG 1863
Db 1816 CTGCTTAACCTTGTGTGTGTGTGTGTGACATGCAATGCCATCCATTTAAGACAGAGAG 1875
Qy 1864 AAGGCTGAATGAAGAAATGTTTGAAGCTGATGACAAATTCATTCAGGCGCAATTC 1923
Db 1876 CAGGCTGAGTTCAAGAAATGTTTGAACCTTATGAGCAATACATTCGAATGCGCATC 1935
Qy 1924 AGATGATATCATCTCAATGAACAGAAATCCGAATTTGAATTTTCCGATACATTTGC 1983
Db 1936 CGGTGATCTCCGCGAGAGAAACCGTGTCCGAACGTAAGCTTACCTGTATATCTGC 1995
Qy 1984 GACAGGAAGTGTCTTTGACAGCCTGCAATTTGATGAACCTTTGATTTGACAGTTGTG 2043
Db 1996 GACACCAAGGCTCTTTTGTGACGCGCTTTCTAGAGAGATTCGGTCTCACTGTGTGTT 2055
Qy 2044 GAGGCAATGACTTGCAGTTTGCACAACTTGCACCTGTATACGCTGAGCAACGCGGAT 2103
Db 2056 GAGTCCATGACCTGTGTGTCTCCGACATTTGCAACCGCTATGCTGTCCAGCTGAGATC 2115
Qy 2104 ATTGTCAATGGAATTTGTGTTTCAATGATTCCTTACATGTGATCAAGCTGTGAC 2163
Db 2116 ATGTTGAACGAGTGTCTGCTTTCACATTTGACCAATGAGGTGCAAGGCTCTCGCG 2175
Qy 2164 ATACTGTGATTTCTTTGAAAGTGTGAAGAAATCATCTGACGTGATTAATCTCC 2223
Db 2176 CTGCTGTGTGATTTCTTTGAAAGTGTGCAAGAGACCCAGCCTGTGACCAAGATCTCG 2235
```

QY	2222	AAAGAGGCTTGAACGAATGAGAGAGATTAATCAATGGAATTTACTCTGGAGAGACTA	2222		
Db	2236	CAGGCGGCTTCTACGCGTATCGAGGAATATCACTGGAAGCTTACTCTGAGAGGCTG	2299		
QY	2284	TTGACCCCTGACAGAGATGATGATTTCTGGAAGCATGTTTCCAACTTGAAAGCCCTGAG	2344		
Db	2296	ATGACCCCTCACCGGCGTCTACGGAATTTGGAAGTACGTCTGAACTTGAGAGGCGTGA	2355		
QY	2344	AGTGTGCTTACCTTGAATGTTTATGCTCTTAAAGTACCGTAAGCTGGCTGATCAATT	2400		
Db	2356	ACCGGCGGCTTACCTTGAAGATGCTGATGCTTCAAGTACCGAGCCATGCTGACACCGTG	2411		
QY	2404	CCATTGGCAGAGAG	2418		
Db	2416	CCGTTGGCGTGGAG	2430		
RESULT 15					
ID	ADC08208	ADC08208 standard; DNA; 2451 BP.			
AC	ADC08208;				
DT	18-DEC-2003	(first entry)			
XX					
DE	Rice DNA sequence Seq ID513 related to grain filling.				
KX					
KW	plant biotechnology; carbohydrate synthesis; carbohydrate metabolism;				
KM	carbohydrate degradation; carbohydrate; plant grain; grain filling; corn;				
KM	tomato; banana; canola; cotton; peanut; sorghum; tobacco; sugarbeet;				
KM	wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain;				
KM	gene; ds; plant.				
XX					
OS	Oryza sativa.				
PN	WO200300905-A2.				
PD	03-JAN-2003.				
XX					
PF	21-JUN-2002; 2002WO-IB002450.				
PR	22-JUN-2001; 2001US-0300112P.				
PR	26-SEP-2001; 2001US-0325277P.				
PR	20-DEC-2001; 2001US-0342327P.				
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.				
XX					
PI	Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;				
PI	Glazebrook J, Katagiri F, Kreps J, Provart N, Rike D;				
XX					
DR	WPI: 2003-229341/22.				
XX	P-PSDB: ADC08209.				
PT	New plant genes encoding polypeptides having an activity involved in or				
PT	associated with the synthesis, metabolism or degradation of carbohydrates				
PT	in the plant grain useful in generating plants having improved				
PT	nutritional properties.				
XX					
PS	Claim 35; SEQ ID NO 513; 130bp; English.				
CC	This invention, in the area of plant biotechnology, relates to novel				
CC	polynucleotides comprising a nucleotide sequence encoding a protein which				
CC	is involved in or associated with the synthesis, metabolism or				
CC	degradation of carbohydrates in the plant grain and the expression of				
CC	which is up-regulated during grain filling. The plant is selected from				
CC	corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,				
CC	sugarbeet, wheat, and rice. The invention may be useful for the				
CC	improvement of protein, oil, starch, fibre and moisture content of the				
CC	cereal grains. In addition, carbohydrate levels may be modified to a more				
CC	desirable level using the present invention. The present sequence is a				
CC	DNA sequence encoding a rice protein of the invention. Note: The sequence				
CC	data for this patent did not form part of the printed specification, but				

Query Match	49.1%; Score 1287.6; DB 9; Length 2451;
Best Local Similarity	71.4%; Pred. No. 0;
Matches 1724; Conservative	0; Mismatches 685; Indels 6; Gaps 2;
Seq Sequence 2451 BP; 557 A; 674 C; 628 G; 592 T; 0 U; 0 Other;	
QY 7 GAGGTCCTCTCACTCGCGTCCACAGTCTCCGTAGCGTTGGATGAGACCTCTTCTGCT	66
Db 19 GACCGCGCTCCAGAGCCGCTCCACAGAGGTAGAGGAGCGCATCGCGCATCTCCCTCCGCGC	78
QY 67 CACAGAAACGAGATTTTGGCCCTTGCTCTCAAGATTCGAGGGCAAAAGGAAATTCGTG	126
Db 79 CACCCCAATAGAGCTCGTCGCGCTCTTACCAAGCGTGTGTAACTTCGAGAAAGGAGATGCTT	138
QY 127 CAACACCATCAAAATTTATTTAGAGT---TTGAAGCTATCCCTGAAGAGAAACAGAAAGAG	183
Db 139 GAGGCCACCAAGATCATTTGCTGAGTACAAACAAGCAATCTCTGAGCGTGAATCGAGAGAG	198
QY 184 CTCGCTAATGAGTCATTTTGAAGTATGAAGGCTAGTCAGAGAGCGATGCTGCT	243
Db 199 CTGAAGACGGTGTCTTTGAGAGATGCTCTGAGAGTGCACAGAGGGAGATGTATCTCC	258
QY 244 CCATGGGTGACATTGCTGTTCCTCAAGGCTGTGTGGAGTACATTAGATGAT	303
Db 259 CCATGGGTGCTCTTGCCATCCCGCCAGGCGCGGTGTCTGGAGGTATGTGAGGGTCAAT	318
QY 304 GTTCAGCGCCCTGTGTTGTTGAGGAATCTGACTGTGCTGATATCTCCACTTCAAGAGAGAG	363
Db 319 GTGAGCAAGCTCGCTGTGAGTGTCTGACAGTCCCGAGTACTTGACGTTCAAGAGAGAG	378
QY 364 CTGTGTGATGAGAAATTCAAATGAGAACTTTGTTGAAATGGAATTTGAGCCCTTCAAC	423
Db 379 CTGTGTGAGAGAGGACCAACCAACAATTTGTGCTTGAAGTTCGATTCGAGCATTCAT	438
QY 424 TCATCATTTCCCGCCCACTCTTCAAAATCCATTGGTATGGTGTGAGATTCCTAAT	483
Db 439 GCCCTCTTCCCTCTCTCTCTGTCTGACAGTCCATTGGCATGTGTGCAATTCCTCAAC	498
QY 484 CGTCACTTTCCGGAATAATGTTTCATGACAGAGAGAGCATGCCCTTGGCTCGAATTC	543
Db 499 AGGCATCTGTATGAAAGCTCTTTCATGACAGAGAGAGCATGACCACTGCTCAATTC	558
QY 544 CTCAGATGTCATTTGTCAAGAGGCGCAAGACATGATGTTGATGACAGAAATTCAGAACTTG	603
Db 559 CTGTGTGACACAACTTCAAGGGCATGACCATGATGTTGAATGACAGAAATTCGAGTCTC	618
QY 604 AATGCTCTTCAACATGTTTGAAGAAACAGAGAGATATCTTGGTACCCTTACCTCTGAG	663
Db 619 AGTCTCTCCAAAGGTGCTGTGAGAGAAAGCTGAGAGATCTTTCTGTCTTTCAGCAGAC	678
QY 664 AACACATGTCCGAATTCGAACACCGGTTCCAGAAATTCGGTTTGAAGAGGTTGGGCT	723
Db 679 ACCCTTACTAGAAATTCACCAAGAGTTCAGGAACCTTGCTGTGGAAGAGGTTGGGCT	738
QY 724 GACACCGCAGAGACGCTGCTGAGATGATCCAACTCTTTTGGATCTTCTTGAAGCAACT	783
Db 739 GACTGTGCAAGGCTTCCAGAGACCATTCACCTCTTTTGAACCTTCTTGAAGCCCT	798
QY 784 GATCTTTCACCCCTTGAAGAGTTCTTTGGAGAAATCCCATGCTTCAATGTTGTGATTT	843
Db 799 GATCGTTCACCCCTCGAAGAGTTCTTTGAACCAATCCCAATGAGGTTCAATGTTGTCAAT	858
QY 844 CTCACCTCCCAACGATATCTTCGCTCAAGACAAATTTTGGGTATCCCGACACGGGTGGG	903
Db 859 ATGTGCCCAACGGTATCTTCGCGCAAGCCAAATGTGGGTATCCCTGAACCTGAGAGGG	918
QY 904 CAGGTTGTTTACATCTTGGATCAAGTCCGAGCTTTGAGAAATGATGATCTCTCCGTATA	963
Db 919 CAGGTTGTTTACATTTTGGACCAAGTCCGTGTATGAGAAATGAGATGCTTCTTGAAGATC	978

QY 964 AAGCAACAAGAGCTCAACATGACCCCTGGATTCCTATTATTACTAGACTTCTCTGAT 1023
Db 979 AAGAGAGAGAGGTCTCAATATTAACACAGCATCCCTTATGTGCACAGAGTGTCTCCCTGAT 1038
QY 1024 GCTGTGGAACAACATGCGGTCAACGACTTGAAGAAAGTATACGGAACAGACACTCGGAT 1083
Db 1039 GCAACTGGACACACTGTGGGTCAAGGTCTTGAAGAGGTCTTGGCACCGACACACCCAC 1098
QY 1084 ATTCTTCAGTACCTCTCAGAACAGAAAAGGAATTTGTTGAAAATGAGATCTCAAGATT 1143
Db 1099 ATCTCACTGTGCTTTTCAAGAACTGAATGGAATTTGTTCCAAAGTGGATCTCAGTTTT 1158
QY 1144 GAAAAAGCTGGCCATTAATTGGAAACCTACACAGAGATGTTGCTCATGAATCTCCAA 1203
Db 1159 G---AGGCTGGCGGTACTGGAACATTCACAGTAGATGTGGGCACGAAATTTGCTGGA 1215
QY 1204 GAGTTGACGGGACGCCGAGATCTGATCATCGGAAACNACAGCGGCAATATGCTGACC 1263
Db 1216 GAGCTCCAGGCCAATCTTGACTTGATCATCGGAACCTACAGTGAAGGAAACCTTGTGCA 1275
QY 1264 TCCTTGTGCGACATAAATTAAGTGTCAACAGTGCACACATGCGCCATGCTTTGAAGA 1323
Db 1276 TGCTTGTGTGACACAAGATGGGTGTACCATTTGATCATTTGCCCATGGCCTTGAAGAA 1335
QY 1324 ACAAAATATCCAGATTCAGATATCTATTGGAAGAACCTTGAAGACAAATACATTTCTCT 1383
Db 1336 ACCAAGTACCCCACTGACCTCTACCTGAAGAAAGTTGAGATTCATACACTTCTCA 1395
QY 1384 TGCCAAATTAAGCTGATCTTTTTCATGAACCACTACAGATTCATCATCAGCACTACT 1443
Db 1396 TGCCAGTTCAAACTGACTTATGTCGATGAACCAAGCTGACTTCTCATATACCAAGTACC 1455
QY 1444 TTCCAGAAATTTGAGAGAGAGACACTGTGTCTAATAAGAGCCACACTGCTTTC 1503
Db 1456 TTCCAAAGAGATTTGCCGGAACAAGAGACACTGTGTGTCAGTATGAGTCTCACATGGCAATTC 1515
QY 1504 ACTCTTCTGGTCTCTAACCGTGTGTATGATGATCATGATGTTTGAATCCCAATTCAC 1563
Db 1516 ACCATGTGCTGTGTACCGGTGTGTCCAGGTATGATGTGTTTGAACCCCAAGTTTAAAC 1575
QY 1564 ATTGTTTCCCTGGTGTGATATGAGATATATCTCCCTTACACCGAAAGAGAGCGGAG 1623
Db 1576 ATTGTCTGCTGGTGTGGGACATGTCCATCTTCTTACTCTGAGTCTCGTAAGAG 1635
QY 1624 TTGAAGCATTTCCATCTGAGATGAGAACCTTCTTTACACCAAGTTGAGATGAAGAA 1683
Db 1636 CTCACTCCCTCCACCCGAGATTTGAGAGTGTGCTTACAGGAAAGTTGACACACAGAG 1695
QY 1684 CACTTATGTGTCTCATAGACCGCAACAGCCATTTCTTTTCAACAATGCCAAGGCTTGAT 1743
Db 1696 CACAAGTTTATGTCTGAGAGAGACAGAAACAGCCATCTTCTCATGTGCTCGTCTTGAC 1755
QY 1744 CGTGTCAAGAACTTAACCGGACCTGTCGAGTGTGCGGCAAGAACCCAAAGTTGCGTGAG 1803
Db 1756 CGTGTCAAGAACTTGACTGTGTGTGTGATGCTGTATGTGCGCAACCTCGCCTGCAAGAG 1815
QY 1804 TTGCTTAACCTCTGATGTGTGTGTGTATGCGCAAGAAAGATCTAAAGATTTGAAAG 1863
Db 1816 CTGGTTAACCTTGTGTGTGTGTGTGTGATGCAATGGCAACCATCTTAAGGACAGAGAG 1875
QY 1864 AAGCTGAATGAAGAAATTTTGAAGTGTATGCAAGTACAACTTGAACGGCCATTC 1923
Db 1876 CAGGCTGAGTTCAAGAAAGATTTTGAACCTTATTTGAGCAATCAACTTGAATGGCCACATC 1935
QY 1924 AGATGATATCATCTCAATGAACAGATTCGAAATGTTGAACCTTTACCGATACATTTGC 1983
Db 1936 CGCTGATCTCCGCGACATGAACCGTGTCCGCAACGGTGAAGCTCTACCGTTACATCTGC 1995
QY 1984 GACACGAAGAGTGCCTTTGTACAGCTGCAATGTATGAAGCCTTTGATGACAGTTGTG 2043
Db 1996 GACACGAAGAGT 2055
QY 2044 GAGGCAATGACTTGGGTTTGGCAACATTCGCACTGTAAAGGTGACACGCGAGATT 2103

Db 2056 GAGTTCATGACCTGTGTCTTCCGCAATTCGCAACGCCCTATGTGTCCAGCTGAATC 2115
QY 2104 ATTGTCCATGGAAATCTGTTTCAACATTTGATCTTTACATGTGTATCAAGCTGTGAC 2163
Db 2116 ATGTTGAACGAGATGTGTGCTTCCACATTTGACCCATACCAAGGTGTACAGGCTCGGCG 2175
QY 2164 ATACTGTGATTTCTTTGAAAAGTGTAAAGAAATCATCTCACTGGATTAAGATCTCC 2223
Db 2176 CTGCTGTGTAGTTCCTTTGAAAGTGTCCAGAAAGACCCAGCCACTGGAACCAAGATCTCG 2235
QY 2224 CAAGAGGCTTTGAACGAATAGAGAGATTAATCATGAAAGTTTACTCGAGAGACTA 2283
Db 2236 CAGGCGGTCTTCAAGCTATGAGAGAAATPACCTGAAAGCTTACTGTGAAGGCTG 2295
QY 2284 TTGACCTGACAGAGTGTATGATTTCTGGAAGCATGTTTCAACCTTGAACGCCGTGAG 2343
Db 2296 ATGACCTTCAACGCGCTGTACGGAATCTGGAAGTACGTCTGAACTGGAGAGGCTGAA 2355
QY 2344 AGTGTGTTTACCTTGAAGATGTTTATGTCTTTAAGTACCGTAAAGCTGCTGAATCATT 2403
Db 2356 ACCGCGCTTACCTTGAAGATGTGTATGCCCTCAAGTACCGCACATGGCTAGCACCGTTC 2415
QY 2404 CCAATGCGAAGAG 2418
Db 2416 CCGTTGCGGTGAG 2430

Search completed: June 5, 2004, 01:13:49
Job time : 1477 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 14:48:19 ; Search time 51 Seconds
(without alignments)
4986.425 Million cell updates/sec

Title: US-10-003-405-2

Perfect score: 4247

Sequence: 1 MAERALTRVHSLRLDELTL.....EMFALKYKLAESVPLAER 806

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	ID	Description
1	4245	100.0	806	10	Q9XGB7 gossypium h
2	3839.5	90.4	805	10	Q9SLS2 Q9SLS2 citrus unsh
3	3815.5	89.8	805	10	Q9SLY1 Q9SLY1 citrus unsh
4	3674.5	86.5	806	10	Q9TOM9 Q9TOM9 pisum sativ
5	3673.5	86.5	806	10	Q9SBL8 Q9SBL8 citrullus l
6	3662.5	86.2	806	10	Q9SLT0 Q9SLT0 pisum sativ
7	3649.5	85.9	805	10	Q8GTJ3 Q8GTJ3 phaseolus v
8	3646.5	85.9	805	10	Q9TOM6 Q9TOM6 medicago tr
9	3637.5	85.6	805	10	Q9XG65 Q9XG65 medicago tr
10	3614.5	85.1	808	10	Q9LXJ5 Q9LXJ5 arabidopsis
11	3536.5	83.3	804	10	Q9AVR8 Q9AVR8 pisum sativ
12	3498.5	82.3	805	10	Q82633 Q82633 lycopersico
13	3488.5	82.1	805	10	Q7Y078 Q7Y078 solanum tub
14	3470.5	81.7	805	10	Q84UC3 Q84UC3 solanum tub
15	3452	81.3	803	10	Q9LMB7 Q9LMB7 chenopodium
16	3443.5	81.1	805	10	Q82691 Q82691 lycopersico

17	3296	77.6	816	10	Q8W1W4 Q8W1W4 bambusa old
18	3293.5	77.5	816	10	Q8LJ74 Q8LJ74 x mokara cv
19	3292	77.5	816	10	Q43706 Q43706 zea mays (m
20	3289.5	77.5	808	10	Q8W1W3 Q8W1W3 bambusa old
21	3287.5	77.4	808	10	Q8W1W2 Q8W1W2 bambusa old
22	3272	77.0	816	10	Q8G523 Q8G523 oryza sativ
23	3270.5	77.0	816	10	Q8LJ75 Q8LJ75 oncidium cv
24	3264.5	76.9	802	10	Q9LKR0 Q9LKR0 sacccharum o
25	3209.5	75.6	815	10	Q43223 Q43223 triticum ae
26	3186.5	75.0	808	10	Q82073 Q82073 triticum ae
27	3071	72.3	811	10	Q9SLY2 Q9SLY2 citricum unsh
28	3056	72.0	811	10	Q9SLV8 Q9SLV8 citrus unsh
29	2996	70.5	811	10	Q84T18 Q84T18 solanum tub
30	2972	70.0	809	10	Q8L5H0 Q8L5H0 zea mays (m
31	2962	69.7	809	10	Q9M111 Q9M111 arabidopsis
32	2956	69.6	796	10	Q93W53 Q93W53 zea mays (m
33	2953	69.5	822	10	Q94G60 Q94G60 beta vulgar
34	2940	69.2	812	10	Q9FRX3 Q9FRX3 pyrus pyrif
35	2934	69.1	811	10	Q9ZPC5 Q9ZPC5 crataegus
36	2892	68.1	808	10	Q9SBD5 Q9SBD5 arabidopsis
37	2837	66.8	809	10	Q9ZPC6 Q9ZPC6 crataegus
38	2487.5	58.6	532	10	Q94CC8 Q94CC8 arabidopsis
39	2312.5	54.5	942	10	Q9FXJ2 Q9FXJ2 arabidopsis
40	2188.5	51.5	454	10	Q9ZRC4 Q9ZRC4 gossypium h
41	2150.5	50.6	798	10	Q7XNX5 Q7XNX5 oryza sativ
42	2148.5	50.6	798	10	Q7XXL1 Q7XXL1 oryza sativ
43	2148	50.6	887	10	Q9FHU4 Q9FHU4 arabidopsis
44	2090.5	49.2	794	16	Q820M5 Q820M5 nicotiana t
45	1821.5	42.9	422	10	Q8W402 Q8W402 nicotiana t

ALIGNMENTS

RESULT 1
Q9XGB7 ID Q9XGB7 PRELIMINARY: PRT: 806 AA.

AC Q9XGB7; 01-NOV-1999 (TRENBLREL. 12, Created)
DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE Sucrose synthase.
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX NCBI_TaxID=3635;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Acacia SJ-2; TISSUE=Fiber;
RA Perez-Grau L., Delmer D.;
RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U73588; AAD28641.1; -;
DR GO: GO:0005058; P:biogenesis; IEA.
DR GO: GO:0005985; P:sucrose metabolism; IEA.
DR InterPro: IPR001296; Glyco_transf.1.
DR InterPro: IPR000368; Sucrose synth.
DR Pfam: PF00862; Sucrose synth.1.
SQ SEQUENCE 806 AA; 92754 MW; 5916FA6B7E0DABEC CRC64;

Query Match 100.0%; Score 4245; DB 10; Length 806;
Best Local Similarity 100.0%; Pred. No. 1.3e-302; Indels 0; Gaps 0;
Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAERALTRVHSLRLDELTLAHRNEITALLSRLEGKGILOHQTILPEALPEERK 60
DB 1 MAERALTRVHSLRLDELTLAHRNEITALLSRLEGKGILOHQTILPEALPEERK 60
QY 61 KLANGAFPEVILKASQEAIVLPWVALAVRPPGVMEYIRVVNVALVVEELTVAEYLHKE 120
DB 61 KLANGAFPEVILKASQEAIVLPWVALAVRPPGVMEYIRVVNVALVVEELTVAEYLHKE 120

```

QY 121 ELVDGSSNGNFVLELDEFPFNSSPFRPTLSKISGVGEFLNRHSAXLPHDKESMHPLE 180
DB 121 ELVDGSSNGNFVLELDEFPFNSSPFRPTLSKISGVGEFLNRHSAXLPHDKESMHPLE 180
QY 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEEYIGTLPPETPCAEFEHRFOEIGLERGW 240
DB 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEEYIGTLPPETPCAEFEHRFOEIGLERGW 240
QY 241 GDTAERVLMIQLLDLLEATDPCTLEKFLGRIPMVFNVLITPHGYFAQDNVLYGYPDTG 300
DB 241 GDTAERVLMIQLLDLLEATDPCTLEKFLGRIPMVFNVLITPHGYFAQDNVLYGYPDTG 300
QY 301 GQVYIILDQVRLAENEMLRIRKQGLNTTPRILITRLLPDAVGTTCQRLKXYGTHS 360
DB 301 GQVYIILDQVRLAENEMLRIRKQGLNTTPRILITRLLPDAVGTTCQRLKXYGTHS 360
QY 361 DILRVPTTEKGIYRKWISREPKWPILETETEDVAHISKEHNTPDLLIGNSDGNIV 420
DB 361 DILRVPTTEKGIYRKWISREPKWPILETETEDVAHISKEHNTPDLLIGNSDGNIV 420
QY 421 ASLNAHKLGVTOCTIAHALEKTKYPSDSIYWKLEDKYHSCQFTADLFAMNHTDFITTS 480
DB 421 ASLNAHKLGVTOCTIAHALEKTKYPSDSIYWKLEDKYHSCQFTADLFAMNHTDFITTS 480
QY 481 TFOEIASGKDTVGOYESHHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 540
DB 481 TFOEIASGKDTVGOYESHHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 540
QY 541 RLKHFHPEIBDLTYKYVNEEHLCVLNDNRNRPILFTMRDLRVKNLTGLVEMCGNPLXR 600
DB 541 RLKHFHPEIBDLTYKYVNEEHLCVLNDNRNRPILFTMRDLRVKNLTGLVEMCGNPLXR 600
QY 601 ELANLVVVGDRRKESKDLBEKAEKMKFELIDKYNLNGQFRWISSQNNRIINVELYRYI 660
DB 601 ELANLVVVGDRRKESKDLBEKAEKMKFELIDKYNLNGQFRWISSQNNRIINVELYRYI 660
QY 661 CDTKGAFOPALYEAFGLTVEAMTCGLPTPATCNGPAEIIYHGKSGFNIDPYHGDOAA 720
DB 661 CDTKGAFOPALYEAFGLTVEAMTCGLPTPATCNGPAEIIYHGKSGFNIDPYHGDOAA 720
QY 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTKWISQRLLITLGVYGFPMKHVSNLERR 780
DB 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTKWISQRLLITLGVYGFPMKHVSNLERR 780
QY 781 ESRRYLEMFVALKYRKLAEVPLAE 806
DB 781 ESRRYLEMFVALKYRKLAEVPLAE 806

RESULT 2
Q9SL52 PRELIMINARY; PRT; 805 AA.
ID 09SL52
AC 09SL52:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Sucrose synthase.
GN CITSUS1-2.
OS Citrus unshiu (Satsuma orange).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=55188;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Miyagawa-wase;
RA Komatsu A.;
RT "Cloning of sucrose synthase (CITSUS1) gene from Citrus.";
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB029401; BAA89049.1; -.
DR GO; GO:0009058; P-biosynthesis; IEA.
DR GO; GO:0005985; P-sucrose metabolism; IEA.
DR InterPro; IPR001296; Glyco_trans_1.

```

```

DR InterPro; IPR000368; Sucrose_synth.
DR Pfam; PF00534; Glycos_transf_1; 1.
DR Pfam; PF00862; Sucrose_synth; 1.
SQ SEQUENCE 805 AA; 92172 MW; 2098FC1F524C266E CRC64;

Query Match 90.4%; Score 3839.5; DB 10; Length 805;
Best Local Similarity 90.7%; Pred. No. 7.2e-273;
Matches 731; Conservative 34; Mismatches 40; Indels 1; Gaps 1;

QY 1 MAERALRVHSIRLRDETTLAHREILALSRLEGKGLIOHIOHILEFEALPEERK 60
DB 1 MAERALRVHSIRLRDETTLAHREILALSRLEGKGLIOHIOHILEFEALPEERK 60
QY 61 KLANGAEFEVLKASQEAIVLPWVALAVRPRGWEXYIRVNVHVLVEELTVAEYLHKE 120
DB 61 HLTGAFGEVLRAQEAIVLPWVALAVRPRGWEXYIRVNVHVLVEELTVAEYLHKE 120
QY 121 ELVDGSSNGNFVLELDEFPFNSSPFRPTLSKISGVGEFLNRHSAXLPHDKESMHPLE 180
DB 121 ELVDGSSNGNFVLELDEFPFNSSPFRPTLSKISGVGEFLNRHSAXLPHDKESMHPLE 180
QY 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEEYIGTLPPETPCAEFEHRFOEIGLERGW 240
DB 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEEYIGTLPPETPCAEFEHRFOEIGLERGW 240
QY 241 GDTAERVLMIQLLDLLEATDPCTLEKFLGRIPMVFNVLITPHGYFAQDNVLYGYPDTG 300
DB 241 GDTAERVLMIQLLDLLEATDPCTLEKFLGRIPMVFNVLITPHGYFAQDNVLYGYPDTG 300
QY 301 GQVYIILDQVRLAENEMLRIRKQGLNTTPRILITRLLPDAVGTTCQRLKXYGTHS 360
DB 301 GQVYIILDQVRLAENEMLRIRKQGLNTTPRILITRLLPDAVGTTCQRLKXYGTHS 360
QY 361 DILRVPTTEKGIYRKWISREPKWPILETETEDVAHISKEHNTPDLLIGNSDGNIV 420
DB 361 DILRVPTTEKGIYRKWISREPKWPILETETEDVAHISKEHNTPDLLIGNSDGNIV 420
QY 421 ASLNAHKLGVTOCTIAHALEKTKYPSDSIYWKLEDKYHSCQFTADLFAMNHTDFITTS 480
DB 421 ASLNAHKLGVTOCTIAHALEKTKYPSDSIYWKLEDKYHSCQFTADLFAMNHTDFITTS 480
QY 481 TFOEIASGKDTVGOYESHHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 540
DB 481 TFOEIASGKDTVGOYESHHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 540
QY 541 RLKHFHPEIBDLTYKYVNEEHLCVLNDNRNRPILFTMRDLRVKNLTGLVEMCGNPLXR 600
DB 541 RLKHFHPEIBDLTYKYVNEEHLCVLNDNRNRPILFTMRDLRVKNLTGLVEMCGNPLXR 600
QY 601 ELANLVVVGDRRKESKDLBEKAEKMKFELIDKYNLNGQFRWISSQNNRIINVELYRYI 660
DB 601 ELANLVVVGDRRKESKDLBEKAEKMKFELIDKYNLNGQFRWISSQNNRIINVELYRYI 660
QY 660 ELVNLVVVGDRRKESKDLBEKAEKMKFELIDKYNLNGQFRWISSQNNRIINVELYRYI 659
DB 660 ELVNLVVVGDRRKESKDLBEKAEKMKFELIDKYNLNGQFRWISSQNNRIINVELYRYI 659
QY 661 CDTKGAFOPALYEAFGLTVEAMTCGLPTPATCNGPAEIIYHGKSGFNIDPYHGDOAA 720
DB 661 CDTKGAFOPALYEAFGLTVEAMTCGLPTPATCNGPAEIIYHGKSGFNIDPYHGDOAA 720
QY 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTKWISQRLLITLGVYGFPMKHVSNLERR 780
DB 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTKWISQRLLITLGVYGFPMKHVSNLERR 780
QY 781 ESRRYLEMFVALKYRKLAEVPLAE 806
DB 781 ESRRYLEMFVALKYRKLAEVPLAE 806

RESULT 3
Q9SLY1 PRELIMINARY; PRT; 805 AA.
ID 09SLY1
AC 09SLY1:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)

```

DE Sucrose synthase.
 GN CITSUS1.
 OS Citrus unshiu (Satsuma orange).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eustosids I; Sapindales; Rutaceae; Citrus.
 OX NCBI_TaxID=55188;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Miyagawa-Wase; TISSUE=fruit;
 RA Komatsu A.;
 RL "Isolation of three sucrose synthase isoforms from Citrus."
 RT Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB022092; BA089905.1; -
 DR GO: GO:0009058; P:biogenesis; IEA.
 DR GO: GO:0005985; P:sucrose metabolism; IEA.
 DR InterPro: IPR001296; Glyco_trans_1.
 DR InterPro: IPR000368; Sucrose_synth.
 DR Pfam: PF00534; Glycos_transf_1; 1.
 DR Pfam: PF00862; Sucrose_synth_1.
 DR SEQUENCE 805 AA; 9211 MW; E2204B24E608328 CRC64;
 SO

Query Match 89.8%; Score 3815.5; DB 10; Length 805;
 Best Local Similarity 90.3%; Pred. No. 4,2e-271;
 Matches 728; Conservative 35; Mismatches 42; Indels 1; Gaps 1;

QY 1 MAERALTIVHSIRRELDITLAAHREIILALSRIGKKGKGILOHNOIILEFPAIPEENRK 60
 DB 1 MAERALTIVHSIRRELDITLAAHREIILALSRIGKKGKGILOHNOIILEFPAIPEENRK 60
 QY 61 KLANGAFEVKASQGEAIVLPWVALAVPRPGWVEYIRVNVHVALVEELTVAEYLAFKE 120
 DB 61 HLTGCAFGEVLRATGEAIVLPWVALAVPRPGWVEYIRVNVHVALVEELTVAEYLAFKE 120
 QY 121 ELVDGSSNGNFVLELDFEPFNSSPPRPTLSKISNGVEFLNRHLSAKLFHDKESHAPLLE 180
 DB 121 ELVDGSSNGNFVLELDFEPFNSSPPRPTLSKISNGVEFLNRHLSAKLFHDKESHAPLLE 180
 QY 121 ELVDGSSNGNFVLELDFEPFNSSPPRPTLSKISNGVEFLNRHLSAKLFHDKESHAPLLE 180
 DB 121 ELVDGSSNGNFVLELDFEPFNSSPPRPTLSKISNGVEFLNRHLSAKLFHDKESHAPLLE 180
 QY 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAFEFHROEIGLENGW 240
 DB 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAFEFHROEIGLENGW 240
 QY 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAFEFHROEIGLENGW 240
 DB 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAFEFHROEIGLENGW 240
 QY 241 GDTAERVLVEMIQLLDLLEATDPTCLEKFLGRIPWVFNVVILTPHGYPAODVNLGYPDTG 300
 DB 241 GDTAERVLVEMIQLLDLLEATDPTCLEKFLGRIPWVFNVVILTPHGYPAODVNLGYPDTG 300
 QY 241 GDTAERVLVEMIQLLDLLEATDPTCLEKFLGRIPWVFNVVILTPHGYPAODVNLGYPDTG 300
 DB 241 GDTAERVLVEMIQLLDLLEATDPTCLEKFLGRIPWVFNVVILTPHGYPAODVNLGYPDTG 300
 QY 301 GOVVYIILDVVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCGRLKXVGTGHS 360
 DB 301 GOVVYIILDVVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCGRLKXVGTGHS 360
 QY 301 GOVVYIILDVVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCGRLKXVGTGHS 360
 DB 301 GOVVYIILDVVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCGRLKXVGTGHS 360
 QY 361 DILRVPTTEKGIYKVMISREKWPYLETYTEVVAHEISKELHGTPLDIIIGNSDGNIV 420
 DB 361 DILRVPTTEKGIYKVMISREKWPYLETYTEVVAHEISKELHGTPLDIIIGNSDGNIV 420
 QY 361 DILRVPTTEKGIYKVMISREKWPYLETYTEVVAHEISKELHGTPLDIIIGNSDGNIV 420
 DB 361 DILRVPTTEKGIYKVMISREKWPYLETYTEVVAHEISKELHGTPLDIIIGNSDGNIV 420
 QY 421 ASLAAHKLGVTQCTAAHLEKTKYPSDSIYWKQLDCKHFSQCFADILANMHTDEIITS 480
 DB 421 ASLAAHKLGVTQCTAAHLEKTKYPSDSIYWKQLDCKHFSQCFADILANMHTDEIITS 480
 QY 421 ASLAAHKLGVTQCTAAHLEKTKYPSDSIYWKQLDCKHFSQCFADILANMHTDEIITS 480
 DB 421 ASLAAHKLGVTQCTAAHLEKTKYPSDSIYWKQLDCKHFSQCFADILANMHTDEIITS 480
 QY 481 TFOEIASGKDVYGOYESHATAFTLPGLYVNVGIDVDFPKFNVISPGADMEIYFPTEEKR 540
 DB 481 TFOEIASGKDVYGOYESHATAFTLPGLYVNVGIDVDFPKFNVISPGADMEIYFPTEEKR 540
 QY 481 TFOEIASGKDVYGOYESHATAFTLPGLYVNVGIDVDFPKFNVISPGADMEIYFPTEEKR 540
 DB 481 TFOEIASGKDVYGOYESHATAFTLPGLYVNVGIDVDFPKFNVISPGADMEIYFPTEEKR 540
 QY 541 RLKGFHPEIEDLLYTKVENEHLVLDNRNKPILFTMPRLDRVKNLTGLVWCKGNPKLR 600
 DB 541 RLKGFHPEIEDLLYTKVENEHLVLDNRNKPILFTMPRLDRVKNLTGLVWCKGNPKLR 600
 QY 541 RLKGFHPEIEDLLYTKVENEHLVLDNRNKPILFTMPRLDRVKNLTGLVWCKGNPKLR 600
 DB 541 RLKGFHPEIEDLLYTKVENEHLVLDNRNKPILFTMPRLDRVKNLTGLVWCKGNPKLR 600
 QY 600 ELVNVLVVGGDRKESKDLSEKAEKMKMSLIDQNKLGQEFKWSQNNRRANGELYI 659
 DB 600 ELVNVLVVGGDRKESKDLSEKAEKMKMSLIDQNKLGQEFKWSQNNRRANGELYI 659
 QY 661 CDTKGAFVOPALYEAFTLVVEAMTCGLPTFATCGGPAEIIYHSGSFNIDPYHGDQA 720
 DB 661 CDTKGAFVOPALYEAFTLVVEAMTCGLPTFATCGGPAEIIYHSGSFNIDPYHGDQA 720
 QY 660 CETKGAFAVOPALYEAFTLVVEAMTCGLPTFATCGGPAEIIYHSGSFNIDPYHGDQA 719

QY 721 DILVDFPEKCKDPSHMDKISQGLKRIEKKYTWKTYSERLTLTGYYGFWMGNSLERR 780
 DB 720 EILDVDFPEKCKDPSHMDKISQGLKRIEKKYTWKTYSERLTLTGYYGFWMGNSLERR 779
 QY 781 ESRRLVEMFVALKYRKLAESVPLAE 806
 DB 780 ESRRLVEMFVALKYRKLAESVPLAE 805

RESULT 4
 ID Q9TOM9 PRELIMINARY; PRT; 806 AA.
 AC Q9TOM9;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Sucrose synthase (EC 2.4.1.14).
 OS Pisum sativum (Garden pea).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eustosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
 OX NCBI_TaxID=3888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=leaf;
 RA Craig J., Barratt P., Tatge H., Dejardin A., Handley L., Gardner C.D.,
 RA Barber L., Wang T., Hedley C., Martin C., Smith A.M.;
 RT "Mutations at the rug4 locus alter the carbon and nitrogen metabolism
 of pea plants through an effect on sucrose synthase";
 RL Plant J. 17:353-362(1999).
 DR EMBL: AJ012080; CA09910.1; -
 DR GO: GO:0046524; F:sucrose-phosphate synthase activity; IEA.
 DR GO: GO:0016757; P:transferase activity, transferring glycosyl. . .; IEA.
 DR GO: GO:0009058; P:biogenesis; IEA.
 DR GO: GO:0005985; P:sucrose metabolism; IEA.
 DR InterPro: IPR001296; Glyco_trans_1.
 DR InterPro: IPR000368; Sucrose_synth.
 DR Pfam: PF00534; Glycos_transf_1; 1.
 DR Pfam: PF00862; Sucrose_synth_1.
 DR Glycoyltransferase; Transferase.
 DR SEQUENCE 806 AA; 92501 MW; CD20A7878C46F76E CRC64;
 SO

Query Match 86.5%; Score 3674.5; DB 10; Length 806;
 Best Local Similarity 86.4%; Pred. No. 9.3e-261;
 Matches 696; Conservative 52; Mismatches 57; Indels 1; Gaps 1;

QY 1 MAERALTIVHSIRRELDITLAAHREIILALSRIGKKGKGILOHNOIILEFPAIPEENRK 60
 DB 1 MAERALTIVHSIRRELDITLAAHREIILALSRIGKKGKGILOHNOIILEFPAIPEENRK 60
 QY 61 KLANGAFEVKASQGEAIVLPWVALAVPRPGWVEYIRVNVHVALVEELTVAEYLAFKE 120
 DB 61 KLANGAFEVKASQGEAIVLPWVALAVPRPGWVEYIRVNVHVALVEELTVAEYLAFKE 120
 QY 61 KLANGAFEVKASQGEAIVLPWVALAVPRPGWVEYIRVNVHVALVEELTVAEYLAFKE 120
 DB 61 KLANGAFEVKASQGEAIVLPWVALAVPRPGWVEYIRVNVHVALVEELTVAEYLAFKE 120
 QY 121 ELVDGSSNGNFVLELDFEPFNSSPPRPTLSKISNGVEFLNRHLSAKLFHDKESHAPLLE 180
 DB 121 ELVDGSSNGNFVLELDFEPFNSSPPRPTLSKISNGVEFLNRHLSAKLFHDKESHAPLLE 180
 QY 121 ELVDGSSNGNFVLELDFEPFNSSPPRPTLSKISNGVEFLNRHLSAKLFHDKESHAPLLE 180
 DB 121 ELVDGSSNGNFVLELDFEPFNSSPPRPTLSKISNGVEFLNRHLSAKLFHDKESHAPLLE 180
 QY 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAFEFHROEIGLENGW 240
 DB 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAFEFHROEIGLENGW 240
 QY 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAFEFHROEIGLENGW 240
 DB 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEYIGTLPETPCAFEFHROEIGLENGW 240
 QY 241 GDTAERVLVEMIQLLDLLEATDPTCLEKFLGRIPWVFNVVILTPHGYPAODVNLGYPDTG 300
 DB 241 GDTAERVLVEMIQLLDLLEATDPTCLEKFLGRIPWVFNVVILTPHGYPAODVNLGYPDTG 300
 QY 241 GDTAERVLVEMIQLLDLLEATDPTCLEKFLGRIPWVFNVVILTPHGYPAODVNLGYPDTG 300
 DB 241 GDTAERVLVEMIQLLDLLEATDPTCLEKFLGRIPWVFNVVILTPHGYPAODVNLGYPDTG 300
 QY 301 GOVVYIILDVVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCGRLKXVGTGHS 360
 DB 301 GOVVYIILDVVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCGRLKXVGTGHS 360
 QY 301 GOVVYIILDVVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCGRLKXVGTGHS 360
 DB 301 GOVVYIILDVVALLENEMILRIKQOGLNTPRILITRLLPDAVGTTCGRLKXVGTGHS 360
 QY 361 DILRVPTTEKGIYKVMISREKWPYLETYTEVVAHEISKELHGTPLDIIIGNSDGNIV 420
 DB 361 DILRVPTTEKGIYKVMISREKWPYLETYTEVVAHEISKELHGTPLDIIIGNSDGNIV 420

DR GO: 0016757; P:transferrase activity, transferring glycosyl. . . , IEA.
DR GO: 0009058; P:biosynthesis, IEA.
DR GO: 0005985, P:sucrose metabolism, IEA.
DR InterPro: IPR001266; Glyco_trans_1.
DR InterPro: IPR000368; Sucrose_synth.
DR Pfam: PF00534; Glycos_transf_1.
DR Pfam: PF00862; Sucrose_synth_1.
DR Glycoyltransferase; Transferrase.
KW SEQUENCE 806 AA; 92413 MW; 7B9B52FF3F0C60D1 CRC64;

Query Match	86.2%;	Score 3662.5;	DB 10;	Length 806;
Best Local Similarity	86.2%;	Pred. No. 7.1e-260;		
Matches 695;	Conservative 50;	Mismatches 60;	Indels 1;	Gaps 1;

QY	1	MAEBALTVHSHREEDJSTLLAHREBEITLALSRIEGKKGILQHOHQIILEEPAIEENRK	60
Db	1	MATDRLTIVHSHREKRLJDETLNANREBILNALLSRIAKGGILQHOQVIAEFEBIEENKQ	60
QY	61	KLANGAFREVLKASQGEATVLPMPVALAVRPREGVEYIRVNVHALVVELTVAEYLHKE	120
Db	61	KLTOGAFGEVLRSTGEATVLPMPVALAVRPRGVEYIRVNVHALVVENTLOAEFLTKFE	120
QY	121	ELVODSSNGNFPVLEDPFEPFNSFPFRPTLSKISGNGVEFARNHLSAKLPHDKESHPLLE	180
Db	121	ELVODSANGNFVLEDPFEPFTASFPFRPTLKSIGNGVOFLNHLISAKLPHDESHAPLLE	180
QY	181	FLRVYCHGKMMULNDRIQNLVALOHVLRKAEYIGTLPETPCAEFEHRFOEIGLERGM	240
Db	181	FLRLHSHYKGTILMNDRIQNPDSLOHVLARKAEYIGTAPVPTPYSEFEHRFOEIGLERGM	240
QY	241	GDTAERVLEMIQLLLDLLEATDPCTLEKFLGRIPMVFNVVILTPHGYPADQNVGLPYDTG	300
Db	241	GDTAERVLIESIQLLDLLEAPDPCTLETFLDRIPMVFNVVILSPHGYPADQNVGLPYDTG	300
QY	301	GOVVYILDOVALENEMILRIKQOGLNTPRLLITRLLPDAVGTTCGRLEKXVGTSHS	360
Db	301	GOVVYILDOVRLSESEMNLRIKQGLDVPRLLITRLLPDAVGTTCGRLEKXVGTSHC	360
QY	361	DILRVPRFTEKGIYAKWKSIRFEKWPVYLETYTEDEVANHESKELAGTTPDLIGNKSDGNIV	420
Db	361	HILRVPRFDQGIYAKWKSIRFE-VMPYLETYTEDEVANHELAKELQCKPDLIGNVSDGNIV	419
QY	421	ASLNAHKLVTOCTTAHALAKTKTYDSDIYMKKLEDKTHFSQCFADLFPAMNHTPFIITS	480
Db	420	ASLNAHKLVTOCTTAHALAKTKTYEESDIYMKKFEKTHFSQCFADLFPAMNHTPFIITS	479
QY	481	TFQBIAGSKDVTGOYESHTAFTLPGLYVRVHGDIVDFPKFNIVSGADMEIYEPYTEERK	540
Db	480	TFQBIAGSKDVTGOYESHTAFTLPGLYVRVHGDIVDFPKFNIVSGADQTIYFPTTESR	539
QY	541	RLKHPHEIEDLLYTKVNEEHLCVYLANDRNKPDILFTMPRLDRVKULTGLVEMCGNPKLR	600
Db	540	RLTSFYPERIEKLLYSTGSENEHICVYKORNKPIIFETMARLDVRVKITGLVEMYGKNAKLR	599
QY	601	ELANLVVVGSPBRKSKDLEEKAEKMKKMFELDKXNLNGOPFWISSOMNRJANVLYRYI	660
Db	600	ELVNLVVVAGDRKSKDLEEKAEKMKKMEHLETKLNGQFWMISSOMNRVANGLERYI	659
QY	661	CDTKGAFVOPALYEAFGLTVVEAMTCGLPTFAITCGSPAEIIVHKSQGENIDPYHGDOQA	720
Db	660	CDTKGAFVOPAYEAFGLTVVEAMTGLPTFATLNGSPAEIIVHKSQGHIDPYHGDRAA	719
QY	721	DILVDFPEKCKDDBSHMDKISGGGLKRIFEKTKTWKISRLLTLTNGVYGFMGHVSNLER	780
Db	720	DILVDFPEKVKDDBSHMDKISGGGLRIIEKTKTWIYSORLLTLTNGVYGFMGHVSNDRL	779
QY	781	ESRRYLEMFYALKYRKLAEVPLAE8806	
Db	780	ESRRILEMFYALKYRKLAEVPLA8805	

ID	ORGANISM	PRELIMINARY	ERT	805 AA
AC	08GTA3			
AC	08GTA3			
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE	Sucrose synthase (EC 2.4.1.13).			
OC	Phaseolus vulgaris (Kidney bean) (French bean).			
OC	Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;			
OC	Spermatophytes, Magnoliophyta, eudicotyledons, core eudicots, rosids;			
OC	eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.			
OC	NCBI_Taxid=3885;			
OC	[1]			
OC	SEQUENCE FROM N.A.			
OC	TISSUE=Root nodules;			
OC	Canas A., Silvente S., Lara M.;			
OC	"Expression of the gene for sucrose synthase during development of			
OC	common bean nodules. "			
OC	Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.			
OC	EMBL, AF15375; AAN76498.1; -.			
OC	GO:GO:0016157; F:sucrose synthase activity; IEA.			
OC	GO:GO:0016757; P:transferase activity, transferring glycosyl. ; IEA.			
OC	GO:GO:0009058; P:biosynthesis; IEA.			
OC	GO:GO:0005985; P:sucrose metabolism; IEA.			
OC	InterPro; IPR001296; Glyco_trans.1.			
OC	InterPro; IPR00368; Sucrose_synth.			
OC	Pfam; PF00534; Glycos_transf.1; 1.			
OC	Pfam; PF00862; Sucrose_synth.1.			
OC	Transferrase; Glycosyltransferase.			
OC	SEQUENCE 805 AA; 92024 MW; 94227641C6AFE863 CRC64;			
Query Match	85.9%; Score 3649.5; DB 10; Length 805;			
Best Local Similarity	86.0%; Pred. No. 6.4e-259;			
Matches 693; Conservative 48; Mismatches 64; Indels 1; Gaps 1;				
QY	1 MAERATPRVHSRLRRDDETLAHHNREITALLSRLEGKXGILHNOHILFEAPLPEENRX 60			
DB	1 MAADRRLRVHSRLRRDDETLAHHNREITALLSRLEAGTGILHNOHQAEEELPEESRQ 60			
QY	61 KLANGAFEEVLAASQEAIVLPPWVALVPRPGWVEYLRVAVHALVVEELTVAEYLHKE 120			
DB	61 KLIDAPBEVLAASQEAIVLPPWVALVPRPGWVEYLRVAVHALVVEELTVAEYLHKE 120			
QY	121 ELVDGSSNGNFVLDEPEPNSSEPRPTLSKISGNGVEFLNRHLSAKLFHDKESMPLLE 180			
DB	121 ELVDGSSNGNFVLDEPEPNSSEPRPTLSKISGNGVEFLNRHLSAKLFHDKESMPLLE 180			
QY	181 FLRVHCHGKRMMLNDRIQINMALQHVLRKAEEVLTGLTPETPCAFEPHRRQELGLEGW 240			
DB	181 FLRVHCHGKRMMLNDRIQINMALQHVLRKAEEVLTGLTPETPCAFEPHRRQELGLEGW 240			
QY	241 GDTARVLEMIQLLDLLEATDPCTLEKFLGRIPMVFNVVLTTHPGYAAQONVAGYPTG 300			
DB	241 GANAEVLESIQLLDLLEADPCTLEKFLGRIPMVFNVVLTTHPGYAAQONVAGYPTG 300			
QY	301 GQVAVYILDQVRALENEMLRIRKQOGLNTPRILITRLLPDAGVTCGQRLKXYGTEHS 360			
DB	301 GQVAVYILDQVRALENEMLRIRKQOGLNTPRILITRLLPDAGVTCGQRLKXYGTEHS 360			
QY	361 DILRVPFTEGIRKATSPREKWPPIETTEDVAHAISELHGTPLIIGNKSQNGIV 420			
DB	361 HILRVPFTEGIRKATSPREKWPPIETTEDVAHAISELHGTPLIIGNKSQNGIV 420			
QY	421 ASLAAHLKGVQCTTAHALLEKTKYPPDSIYVKKLEDKYHFSQCFADLFAANNHDFITS 480			
DB	420 ASLAAHLKGVQCTTAHALLEKTKYPPDSIYVKKLEEKHFSQCFADLFAANNHDFITS 479			
QY	481 TFOELAGSKDVGQYESHATFPLDGLYRVVAGIDVDPKFNIVSPGADMEIYFYTEBK 540			
DB	480 TFOELAGSKDVGQYESHATFPLDGLYRVVAGIDVDPKFNIVSPGADQTIYFSPKETS 539			
QY	541 RLKGFHPETEDLLYTKKNEBEHLCVLNDRAKPIFLTFMRDLRVKVLGLVWCGKNPLR 600			
DB	540 RLTSHPETEDLLYTKKNEBEHLCVLNDRAKPIFLTFMRDLRVKVLGLVWCGKNPLR 599			

QY 601 ELANLVVVGDRRKESKDLBEKAEKMKFELIDKYNNGCFRWSISQNNRIRNVELYRYI 660
 DB 600 ELVNLVVVAGDRRKESKDLBEKAEKMKVSLIETYNKNGCFRWSISQNNRIRNVELYRYI 659
 QY 661 CDTGAFVQPALYEAFLITVTEAMTCGLPTATCNGGPAEIIYVKGSGFNIDPHYGOAA 720
 DB 660 SDTGAFVQPALYEAFLITVTEAMTCGLPTATCNGGPAEIIYVKGSGFNIDPHYGOAA 719
 QY 721 DIIVDFPEKCKDSDHMDKISQGLKRIEKKYTKWISERLLTTLTGYYGFMKHSNLER 780
 DB 720 DIIVDFPEKCKDSDHMDKISQGLKRIEKKYTKWISERLLTTLTGYYGFMKHSNLER 779
 QY 781 ESRRYLEMFPALKYRKLAESVPLAE 806
 DB 780 ESRRYLEMFPALKYRKLAESVPLAE 805

RESULT 8

Q9TOM6 PRELIMINARY; PRT; 805 AA.
 AC Q9TOM6; 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Sucose synthase (EC 2.4.1.13).
 GN SUCS1.
 OS Medicago truncatula (Barrel medic).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
 NCBI_TaxID=3880;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Jemalong; TISSUE=Root nodules;
 RA MEDLINE=99254811; PubMed=10323232;
 RX Hohnjec N., Becker J.D., Puehler A., Perllick A.M., Kuester H.,
 RT "Genomic organisation and expression properties of the MsSUC1 gene,
 RT which encodes a node-enhanced sucrose synthase from the model legume
 RT Medicago truncatula.";
 RL Mol. Gen. Genet. 261:514-522 (1999).
 DR EMBL; AJ131943; CAB40794.1; -;
 DR GO; GO:0016157; F:sucrose synthase activity; IEA.
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR GO; GO:0005985; P:sucrose metabolism; IEA.
 DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR000368; Sucose_synth.
 DR Pfam; PF00534; Glycos_transf_1.
 DR Pfam; PF00862; Sucose_synth; 1.
 KM Glycosyltransferase; Transferase.
 SQ SEQUENCE 805 AA; 92293 MW; 6F1DE955EC3DB1A6 CRC64;

Query Match 85.9%; Score 3646.5; DB 10; Length 805;
 Best Local Similarity 85.6%; Pred. No. 1,le-256;
 Matches 690; Conservative 53; Mismatches 62; Indels 1; Gaps 1;

QY 1 MAERALTIVHSRLERLDETLTAHREBIALSRLEGKGILOHQLFEPAIPEERK 60
 DB 1 MATERLTIVHSRLERLDETLTAHREBIALSRLEGKGILOHQLFEPAIPEERK 60
 QY 61 KLANGAFPEVILKASGEAIVLPPWVALAVPRGWEXIRVNVHVLVEELTVAEYLHPKE 120
 DB 61 KLTDGAFGEVILKASGEAIVLPPWVALAVPRGWEXIRVNVHVLVEELTVAEYLHPKE 120
 QY 121 ELVNGSSNGNRYLLEDFEPNPSFPRPTLSKISNGVFLRHRHSAKLPHKESHPLE 180
 DB 121 ELVNGSSNGNRYLLEDFEPNPSFPRPTLSKISNGVFLRHRHSAKLPHKESHPLE 180
 QY 181 FLRVHCHGKMMMLNDRIQNTALNLOHVARKABEYIGTLPPETPCAEFERFOEIGLERGM 240
 DB 181 FLRLHSYKGTLMMLNDRIQNPDSLOHVRKABEYISTIDPETPYSEFHRFOEIGLERGM 240

QY 241 GDTAERVLEMIQILLDLLEATDPCTLEKFLGRIPMWNVNIITPHGYFAQDNVLGYPTG 300
 DB 241 GDTAERVLESIQILLDLLEAPDDCTLETFIDRIIPMWNVNIITPHGYFAQDNVLGYPTG 300
 QY 301 GQVYVILDOVRALENEMLRIRKQAGNITPRILITLPLDPAVTTGGORLEKYGTEHS 360
 DB 301 GQVYVILDOVRALENSESLIRKQAGNITPRILITLPLDPAVTTGGORLEKYGTEHS 360
 QY 361 DIIAVPPEKGIYRKWISRFPEKWPYLETTYEDVAHESKELAGTDPDLIIGNSSDGNIV 420
 DB 361 HILAVPPEKGIYRKWISRFPEKWPYLETTYEDVAHESKELAGTDPDLIIGNSSDGNIV 419
 QY 421 ASLIAHKLGVTOCTIAHALERKTPDSDIYWKLEDRYHSCOPTADLFAMNHTDFITTS 480
 DB 420 ASLIAHKLGVTOCTIAHALERKTPDSDIYWKLEDRYHSCOPTADLFAMNHTDFITTS 479
 QY 481 TPQEIAGSKDVGYESHTAFTLPGLRVNVHGDVPPKFNIVSPGADMEYFPYTEKR 540
 DB 480 TPQEIAGSKDVGYESHTAFTLPGLRVNVHGDVPPKFNIVSPGADMEYFPYTEKR 539
 QY 541 RLKHFPEIEDLYTKVNEEHLCVLDRNKPILFTMPRLDRVNLTLGVEMCGKNPKLR 600
 DB 540 RLTSFYFEIEELLYSSVNEEHLICVLDRNKPILFTMPRLDRVNLTLGVEMCGKNPKLR 599
 QY 601 ELANLVVVGDRRKESKDLBEKAEKMKFELIDKYNNGCFRWSISQNNRIRNVELYRYI 660
 DB 600 ELVNLVVVAGDRRKESKDLBEKAEKMKVSLIETYNKNGCFRWSISQNNRIRNVELYRYI 659
 QY 661 CDTGAFVQPALYEAFLITVTEAMTCGLPTATCNGGPAEIIYVKGSGFNIDPHYGOAA 720
 DB 660 CDTGAFVQPALYEAFLITVTEAMTCGLPTATCNGGPAEIIYVKGSGFNIDPHYGOAA 719
 QY 721 DIIVDFPEKCKDSDHMDKISQGLKRIEKKYTKWISERLLTTLTGYYGFMKHSNLER 780
 DB 720 DIIVDFPEKCKDSDHMDKISQGLKRIEKKYTKWISERLLTTLTGYYGFMKHSNLER 779
 QY 781 ESRRYLEMFPALKYRKLAESVPLAE 806
 DB 780 ESRRYLEMFPALKYRKLAESVPLAE 805

RESULT 9

Q9XG65 PRELIMINARY; PRT; 805 AA.
 AC Q9XG65; 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
 DE Sucose synthase (EC 2.4.1.13).
 GN SUCS1.
 OS Medicago truncatula (Barrel medic).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
 NCBI_TaxID=3880;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Chori;
 RA MEDLINE=99254811; PubMed=10323232;
 RX Hohnjec N., Becker J.D., Puehler A., Perllick A.M., Kuester H.,
 RT "Genomic organisation and expression properties of the MsSUC1 gene,
 RT which encodes a node-enhanced sucrose synthase from the model legume
 RT Medicago truncatula.";
 RL Mol. Gen. Genet. 261:514-522 (1999).
 DR EMBL; AJ131964; CAB40795.1; -;
 DR GO; GO:0016157; F:sucrose synthase activity; IEA.
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
 DR GO; GO:0009058; P:biosynthesis; IEA.
 DR GO; GO:0005985; P:sucrose metabolism; IEA.
 DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR000368; Sucose_synth.
 DR Pfam; PF00534; Glycos_transf_1.
 DR Pfam; PF00862; Sucose_synth; 1.


```

Db      663 DTKGFVOPALYEAGLTVVEAMTQGLPFATCNGGPAEIIYHGKSGHIDPFYHDDKAAE 722
Qy      722 ILVDFEKKCKDPKSHMDKISOGGLKRIEKKYTWKTYSERLLTLTVGFMKXVSLJEBRE 781
Db      723 SLADFTCKKDPKSHMDKISOGGLKRIEKKYTWKTYSERLLTLTVGFMKXVSLJEBRE 782
Qy      782 SRRYLEMFYALKYRKLAESVPLAEE 806
Db      783 SRRYLEMFYALKYRKLAESVPLAEE 807

```

RESULT 11

```

Q9AVR8      PRELIMINARY;      PRT;      804 AA.
ID 09AVR8
AC 09AVR8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Sucrose synthase isoform 3 (EC 2.4.1.14).
GN SUS3.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eustosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3886;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=leaf;
RA Barratt P., Barber L., Edwards A., Kruger N., Smith A.M., Wang T.,
RA Martin C.;
RT "Multiple Distinct Isoforms of Sucrose Synthase in the Pea Plant.";
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ311496; CAC32462.1; -.
DR GO; GO:0046552; F:sucrose-phosphate synthase activity; IEA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0005985; P:sucrose metabolism; IEA.
DR InterPro; IPR001296; Glyco_trans_1.
DR InterPro; IPR00368; Sucrose_synth.
DR Pfam; PF00534; Glycos_transf_1.
DR Pfam; PF00862; Sucrose_synth; 1.
DR GlycoSyLtransferase; Transferase.
SQ SEQUENCE 804 AA; 91941 MW; 6669B5D4E7A7370DE CRC64;

```

```

Query Match      83.3%; Score 3536.5; DB 10; Length 804;
Best Local Similarity 83.3%; Pred. No. 1.2e-250;
Matches 668; Conservative 66; Mismatches 67; Indels 1; Gaps 1;

```

```

Qy      5 ALTVHSLRERLDETLTAHREIALLSRIEGKGGLIOHQIILEFEAIPENRKLAN 64
Db      3 SLTHSTSLRQRDETLTAHREIALLSRIEGKGGLIOHQIILEFEAIPENRKLAN 62
Qy      65 GAFFEVYKASQBAIYLPVVALAVPRPGVWEYIRVNVHALVVEELTVAEYHKEEELVD 124
Db      63 GYFGEVLSTQBAIYLPVVALAVPRPGVWEYIRVNVHALVVEELTVAEYHKEEELVE 122
Qy      125 GSSNPNFVLELPEPPNSFRPTLSKISGVVEFLNRHLSAKLPHDKSMPLLEPLRV 184
Db      123 GSSNPNFVLELPEPPNSFRPTLSKISGVVEFLNRHLSAKLPHDKSMPLLEPLRV 182
Qy      185 HCHKKKNMMLANDRIQNLNLQHLVLRKAEYLGTLPEETPCAFERHFOEIGLERMGDTA 244
Db      183 HHHNKTIVNDRIQNLNLQHLVLRKAEYLGTLPEETPCAFERHFOEIGLERMGDTA 242
Qy      245 ERYVLEMIQLLDLLEATDPCCTLEKELGRIEMVFNVTYLTPEHYFPQDNVLTGPTDGGQV 304
Db      243 ERYVLEMIQLLDLLEATDPCCTLEKELGRIEMVFNVTYLTPEHYFPQDNVLTGPTDGGQV 302
Qy      305 YLLDQVRLAENEMLRIRKGGGLNTPRIILTRLLPDAGCTCGGRLKGVGTEHSDILR 364
Db      303 YLLDQVRLAENEMLRIRKGGGLNTPRIILTRLLPDAGCTCGGRLKGVGTEHSDILR 362

```

```

Qy      365 VPFRTKEGIYKRWISREPKWPIYLETYTEVDAHESKELHGTPLLIGKXSDGNIVASLL 424
Db      363 VPFRTKEGIYKRWISREPKWPIYLETYTEVDAHESKELHGTPLLIGKXSDGNIVASLL 421
Qy      425 AHKLGVTQCTIAHLEKTKYKPDSDIYWKLEDKHFSCQFADLFPAMNHTPFIITSTOE 484
Db      422 AHKLGVTQCTIAHLEKTKYKPDSDIYWKLEDKHFSCQFADLFPAMNHTPFIITSTOE 481
Qy      485 IAGSKDITVQYESHHTAFTLPGLYRVVHGIDVDFDKFNIVSPGADMEIYFYTEERKRLKH 544
Db      482 IAGSKDITVQYESHHTAFTLPGLYRVVHGIDVDFDKFNIVSPGADMEIYFYTEERKRLKH 541
Qy      545 FHPRIEDILYTKVNEEHLVCLNDRNKPILFTMPRLDRVKVLTGLVEMCGKPKIRELAN 604
Db      542 FHPRIEDILYTKVNEEHLVCLNDRNKPILFTMPRLDRVKVLTGLVEMCGKPKIRELAN 601
Qy      605 LVVVGSRBRKESKOLEEAEKMKMPELIDKYNLNGQPFWISSOMNRJNVLYRIGCTK 664
Db      602 LVVVGSRBRKESKOLEEAEKMKMPELIDKYNLNGQPFWISSOMNRJNVLYRIGCTK 661
Qy      665 GAFVOPALYEAGLTVVEAMTQGLPFATCNGGPAEIIYHGKSGFNIDPYHGDQADILY 724
Db      662 GAFVOPALYEAGLTVVEAMTQGLPFATCNGGPAEIIYHGKSGFNIDPYHGDQADILY 721
Qy      725 DPFKECKKDPKSHMDKISOGGLKRIEKKYTWKTYSERLLTLTVGFMKXVSLJEBRE 784
Db      722 DPFKECKKDPKSHMDKISOGGLKRIEKKYTWKTYSERLLTLTVGFMKXVSLJEBRE 781
Qy      785 YLEMFYALKYRKLAESVPLAEE 806
Db      782 YLEMFYALKYRKLAESVPLAEE 803

```

RESULT 12

```

ID 082693      PRELIMINARY;      PRT;      805 AA.
AC 082693;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Sucrose synthase (EC 2.4.1.13).
GN SUS2.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC Cloning and expression of a second tomato (Lycopersicon esculentum)
RT sucrose synthase gene.
RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ011535; CA09681.1; -.
DR GO; GO:0016157; F:sucrose synthase activity; IEA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0005985; P:sucrose metabolism; IEA.
DR InterPro; IPR001296; Glyco_trans_1.
DR InterPro; IPR00368; Sucrose_synth.
DR Pfam; PF00534; Glycos_transf_1.
DR Pfam; PF00862; Sucrose_synth; 1.
DR GlycoSyLtransferase; Transferase.
SQ SEQUENCE 805 AA; 92465 MW; E6879774C80092E8 CRC64;

```

```

Query Match      82.3%; Score 3494.5; DB 10; Length 805;
Best Local Similarity 81.8%; Pred. No. 1.5e-247;
Matches 659; Conservative 68; Mismatches 78; Indels 1; Gaps 1;

```

```

Qy      1 MAERALTIVHSLRERLDETLTAHREIALLSRIEGKGGLIOHQIILEFEAIPENRKL 60
Db      1 MAERALTIVHSLRERLDETLTAHREIALLSRIEGKGGLIOHQIILEFEAIPENRKL 60

```

QY 61 KLANGAFPEVLKASGEAIVLPWVALAARPPGWWEYIRVNVHVLVVEELTVAEYLHKE 120
 DB 61 KLNHAFPEVLKSTGEAIVLPWVALAARPPGWWEYIRVNVHVLVVEELTVAEYLHKE 120
 QY 121 ELVDSSNGNFVLELDPEFPNSPPRTLSKISGVGEFLNRHLSAKLFHDKESWHPLE 180
 DB 121 ELVDASNGNFVLELDPEFPNSPPRTLSKISGVGEFLNRHLSAKLFHDKESWHPLE 180
 QY 181 FLRVCHGKMMMLNDRIQNLALQHLVLRKAEVYGLTPETPCAEFEHFOEIGLERGM 240
 DB 181 FLRAHHYKGMMLNDRIQNLALQHLVLRKAEVYGLTPETPCAEFEHFOEIGLERGM 240
 QY 241 GPTARVLEMIQLDLLEADPCTLEKFLGRIPWVFNVLTPHGYFAQDNVLYGPTG 300
 DB 241 GPTARVLEMIQLDLLEADPCTLEKFLGRIPWVFNVLTPHGYFAQDNVLYGPTG 300
 QY 301 GOVVYILQVPALEBEMLRKIKEGGLDIPRILITRLLPDAVGTTCGRLEKVGTEHS 360
 DB 301 GOVVYILQVPALEBEMLRKIKEGGLDIPRILITRLLPDAVGTTCGRLEKVGTEHS 360
 QY 361 DILRVPFRTEKGIYVKWISREKWPYLETYEDVAHEISKELHGTPLLIGNSXGDNIV 420
 DB 361 DILRVPFRTEKGIYVKWISREKWPYLETYEDVAHEISKELHGTPLLIGNSXGDNIV 420
 QY 421 ASLHAKLGVTQCTIAHALEKTKYPSDIYWKLEEDKHFSCQFTADLPFAMHPTFIITS 480
 DB 421 ASLHAKLGVTQCTIAHALEKTKYPSDIYWKLEEDKHFSCQFTADLPFAMHPTFIITS 480
 QY 481 TFOEJAGSKDVTGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 540
 DB 481 TFOEJAGSKDVTGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 540
 QY 541 RUKHHPREIDELLYKVENEBHLVYVNDKRPILFTMPRLDRVKNLTGLVWCGNPKLR 600
 DB 541 RUKHHPREIDELLYKVENEBHLVYVNDKRPILFTMPRLDRVKNLTGLVWCGNPKLR 600
 QY 601 ELANLVVVGGRKRSKDLBEAKEMKMPFLIDKYNLNGOFPMISSOMNRIRNVLYRYI 660
 DB 601 ELANLVVVGGRKRSKDLBEAKEMKMPFLIDKYNLNGOFPMISSOMNRIRNVLYRYI 660
 QY 660 ADTKGAFVQAPAYEAFGLTVEAMTCGLPTFATNKGPAEIIIVHSGSFHIDPHYGEQAA 719
 DB 660 ADTKGAFVQAPAYEAFGLTVEAMTCGLPTFATNKGPAEIIIVHSGSFHIDPHYGEQAA 719
 QY 721 DILVDFPEKCKDPBSHMDKISOGGLKRIEKKYTWKISYERLLTLTGTVGFWMGVSNLERR 780
 DB 721 DILVDFPEKCKDPBSHMDKISOGGLKRIEKKYTWKISYERLLTLTGTVGFWMGVSNLERR 780
 QY 780 ESRRLVLEFVALKRYKLAESVPLAE 806
 DB 780 ESRRLVLEFVALKRYKLAESVPLAE 806
 QY 780 ESRRLVLEFVALKRYKLAESVPLAE 806
 DB 780 ESRRLVLEFVALKRYKLAESVPLAE 806

RESULT 13
 QY078 PRELIMINARY; PRT; 805 AA.
 AC QY078; PRT; 805 AA.
 DT 01-OCT-2003 (Tremblrel. 25, Created)
 DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Sucrose synthase 4 (EC 2.4.1.13).
 GN SUS4.
 OS Solanum tuberosum subsp. tuberosum.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC lamiales; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=90692;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Desiree; TISSUE=Tuber;
 RA Baroja-Fernandez E., Munoz F.J., Saikusa T., Rodriguez-Lopez M.,
 RA Akazawa T., Pozueta-Romero J.;
 RT "Sucrose synthase catalyzes the de novo production of ADP-glucose

RT linked to starch biosynthesis in heterotrophic tissues of plants."
 RL Plant Cell Physiol. 44:500-509 (2003).
 DR EMBL: AJ537575; CAD6188.1; -;
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 805 AA; 92443 MW; 7272FD4E5367A521 CRC64;
 Query Match 82.1%; Score 3488.5; DB 10; Length 805;
 Best Local Similarity 81.8%; Pred. No. 4.2e-247;
 Matches 659; Conservative 68; Mismatches 78; Indels 1; Gaps 1;
 QY 1 MAERLVTVHSIRERDELTLAHRNEILLALSRIEKGKGLIQHQLLEFPAIDENRK 60
 DB 1 MAERLVTVHSIRERDELTLAHRNEILLALSRIEKGKGLIQHQLLEFPAIDENRK 60
 QY 61 KLANGAFPEVLKASGEAIVLPWVALAARPPGWWEYIRVNVHVLVVEELTVAEYLHKE 120
 DB 61 KLNHAFPEVLKSTGEAIVLPWVALAARPPGWWEYIRVNVHVLVVEELTVAEYLHKE 120
 QY 121 ELVDSSNGNFVLELDPEFPNSPPRTLSKISGVGEFLNRHLSAKLFHDKESWHPLE 180
 DB 121 ELVDASNGNFVLELDPEFPNSPPRTLSKISGVGEFLNRHLSAKLFHDKESWHPLE 180
 QY 181 FLRVCHGKMMMLNDRIQNLALQHLVLRKAEVYGLTPETPCAEFEHFOEIGLERGM 240
 DB 181 FLRAHHYKGMMLNDRIQNLALQHLVLRKAEVYGLTPETPCAEFEHFOEIGLERGM 240
 QY 241 GPTARVLEMIQLDLLEADPCTLEKFLGRIPWVFNVLTPHGYFAQDNVLYGPTG 300
 DB 241 GPTARVLEMIQLDLLEADPCTLEKFLGRIPWVFNVLTPHGYFAQDNVLYGPTG 300
 QY 301 GOVVYILQVPALEBEMLRKIKEGGLDIPRILITRLLPDAVGTTCGRLEKVGTEHS 360
 DB 301 GOVVYILQVPALEBEMLRKIKEGGLDIPRILITRLLPDAVGTTCGRLEKVGTEHS 360
 QY 361 DILRVPFRTEKGIYVKWISREKWPYLETYEDVAHEISKELHGTPLLIGNSXGDNIV 420
 DB 361 DILRVPFRTEKGIYVKWISREKWPYLETYEDVAHEISKELHGTPLLIGNSXGDNIV 420
 QY 421 ASLHAKLGVTQCTIAHALEKTKYPSDIYWKLEEDKHFSCQFTADLPFAMHPTFIITS 480
 DB 421 ASLHAKLGVTQCTIAHALEKTKYPSDIYWKLEEDKHFSCQFTADLPFAMHPTFIITS 480
 QY 481 TFOEJAGSKDVTGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 540
 DB 481 TFOEJAGSKDVTGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 540
 QY 541 RUKHHPREIDELLYKVENEBHLVYVNDKRPILFTMPRLDRVKNLTGLVWCGNPKLR 600
 DB 541 RUKHHPREIDELLYKVENEBHLVYVNDKRPILFTMPRLDRVKNLTGLVWCGNPKLR 600
 QY 601 ELANLVVVGGRKRSKDLBEAKEMKMPFLIDKYNLNGOFPMISSOMNRIRNVLYRYI 660
 DB 601 ELANLVVVGGRKRSKDLBEAKEMKMPFLIDKYNLNGOFPMISSOMNRIRNVLYRYI 660
 QY 660 ADTKGAFVQAPAYEAFGLTVEAMTCGLPTFATNKGPAEIIIVHSGSFHIDPHYGEQAA 719
 DB 660 ADTKGAFVQAPAYEAFGLTVEAMTCGLPTFATNKGPAEIIIVHSGSFHIDPHYGEQAA 719
 QY 721 DILVDFPEKCKDPBSHMDKISOGGLKRIEKKYTWKISYERLLTLTGTVGFWMGVSNLERR 780
 DB 721 DILVDFPEKCKDPBSHMDKISOGGLKRIEKKYTWKISYERLLTLTGTVGFWMGVSNLERR 780
 QY 780 ESRRLVLEFVALKRYKLAESVPLAE 806
 DB 780 ESRRLVLEFVALKRYKLAESVPLAE 806
 QY 780 ESRRLVLEFVALKRYKLAESVPLAE 806
 DB 780 ESRRLVLEFVALKRYKLAESVPLAE 806

RESULT 14
 QY084UC3 PRELIMINARY; PRT; 805 AA.
 ID QY084UC3; PRT; 805 AA.
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 14:51:10 ; Search time 49 Seconds

(without alignments)
4627.734 Million cell updates/sec

Title: US-10-003-405-2

Perfect score: 4247
Sequence: 1 MAERALTRVHSLRRLDETL.....EMFYALKRYKLAESVPLAE 806

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155919 seqs, 28138677 residues

Total number of hits satisfying chosen parameters: 1155919

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/2/pubpaa/US09 NEW_PUB.pep:*

13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/2/pubpaa/US10 NEW_PUB.pep:*

17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4245	100.0	806	US-10-003-405-2	Sequence 2, Appl1
2	3763.5	88.6	805	US-10-424-599-258404	Sequence 258404, A
3	3744.5	88.2	806	US-10-425-114-4297	Sequence 54297, A
4	3743.5	88.1	805	US-10-424-599-154986	Sequence 154986, A
5	3658.5	86.1	806	US-10-424-599-154301	Sequence 154301, A
6	3655.5	86.1	811	US-10-424-599-154301	Sequence 154301, A
7	3652.5	86.0	806	US-10-424-599-154300	Sequence 154300, A
8	3652.5	86.0	811	US-10-425-114-55184	Sequence 55184, A
9	3643.5	85.8	805	US-10-137-036-77	Sequence 77, Appl1
10	3643.5	85.8	805	US-10-393-840-44	Sequence 44, Appl1
11	3643.5	85.8	805	US-10-393-840-144	Sequence 144, Appl1
12	3299	77.7	823	US-10-425-114-53296	Sequence 53296, A
13	3299	77.7	823	US-10-425-114-5696	Sequence 5696, A
14	3299	77.7	823	US-10-425-114-57139	Sequence 57139, A
15	3299	77.7	823	US-10-425-114-72813	Sequence 72813, A

16	3295	77.6	816	US-10-080-114A-7	Sequence 7, Appl1
17	3272.5	77.1	807	US-10-425-114-57840	Sequence 57840, A
18	3272.5	77.1	811	US-10-425-114-53965	Sequence 53965, A
19	3270.5	77.0	815	US-10-289-757-170	Sequence 170, Appl1
20	3268.5	77.0	802	US-10-080-114A-5	Sequence 5, Appl1
21	3267	76.9	814	US-10-289-757-76	Sequence 76, Appl1
22	3254.5	76.6	808	US-10-289-757-169	Sequence 169, Appl1
23	3250.5	76.5	808	US-10-289-757-74	Sequence 74, Appl1
24	32348	76.5	816	US-10-289-757-77	Sequence 77, Appl1
25	3234.5	76.2	808	US-10-289-757-75	Sequence 75, Appl1
26	3158.5	74.4	670	US-10-425-114-49869	Sequence 49869, A
27	3127.5	73.6	824	US-10-425-114-56196	Sequence 56196, A
28	3005	70.8	824	US-10-425-114-45934	Sequence 45934, A
29	2892	70.4	812	US-10-424-599-241731	Sequence 241731, A
30	2872	70.0	809	US-10-080-114A-12	Sequence 12, Appl1
31	2861	69.7	802	US-10-080-114A-2	Sequence 2, Appl1
32	2914.5	68.6	619	US-10-425-114-68121	Sequence 68121, A
33	2875	67.7	797	US-10-217-939-30	Sequence 30, Appl1
34	2866.5	67.5	805	US-10-217-939-28	Sequence 28, Appl1
35	2826.5	66.6	593	US-10-425-114-50109	Sequence 50109, A
36	2676.5	63.0	652	US-10-425-114-46682	Sequence 46682, A
37	2675.5	63.0	639	US-10-425-114-58450	Sequence 58450, A
38	2552.5	60.1	596	US-10-425-114-62525	Sequence 62525, A
39	2484	58.5	651	US-10-425-114-49911	Sequence 49911, A
40	2318.5	54.6	570	US-10-425-114-54291	Sequence 54291, A
41	2286.5	53.8	527	US-10-425-114-62319	Sequence 62319, A
42	2254	53.1	567	US-10-424-599-241717	Sequence 241717, A
43	2177.5	51.3	511	US-10-425-114-58452	Sequence 58452, A
44	2138.5	50.4	628	US-10-424-599-244151	Sequence 244151, A
45	2137.5	50.3	503	US-10-425-114-72784	Sequence 72784, A

ALIGNMENTS

RESULT 1

US-10-003-405-2

Sequence 2, Application US/10003405

Publicat ion No. US20020116736A1

GENERAL INFORMATION:

APPLICANT: Ruan, Yong-Ling

APPLICANT: Fuhbank, Robert T.

APPLICANT: Danny, Llewellyn J.

TITLE OF INVENTION: Modification of sucrose synthase gene expression in plant tissue

TITLE OF INVENTION: uses thereof

FILE REFERENCE: GHSUSY WO1

CURRENT APPLICATION NUMBER: US/10/003, 405

CURRENT FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: 60/251852

PRIOR FILING DATE: 2000-12-08

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 806

TYPE: PRT

ORGANISM: Gossypium hirsutum

FEATURE:

NAME/KEY: misc feature

LOCATION: (414)..(414)

OTHER INFORMATION: The 'Xaa' at location 414 stands for Asn, Asp, His, or Tyr.

NAME/KEY: misc feature

LOCATION: (1240)..(1240)

OTHER INFORMATION: n = any nucleotide (a,g,c,t)

US-10-003-405-2

Query Match 100.0%; Score 4245; DB 13; Length 806;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAERALTRVHSLRRLDETLAHRNEITLALSRLEGKKGILQHHQIILFEAIPENRK 60

DB 1 MAERALTRVHSLRRLDETLAHRNEITLALSRLEGKKGILQHHQIILFEAIPENRK 60

QY 61 KLANGAFEEVLKASQGEAIVLPWPVALAVRPPGWMEYIRVNVHVLVVEELTVAEYLHPKE 120
DB 61 KLANGAFEEVLKASQGEAIVLPWPVALAVRPPGWMEYIRVNVHVLVVEELTVAEYLHPKE 120
QY 121 ELVNGSSNGNFVLELDEFPNSSFPRLTSGKSGVGFNLRHSAKLPHDKESHPLLE 180
DB 121 ELVNGSSNGNFVLELDEFPNSSFPRLTSGKSGVGFNLRHSAKLPHDKESHPLLE 180
QY 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEEYIGTLPPETPCAEFHRFOEIGLERGW 240
DB 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEEYIGTLPPETPCAEFHRFOEIGLERGW 240
QY 241 GDTAERVLEMIQLLDLLEATDPCLEKFLGRIPMVENVVILTPHGYFAQDNVLYGYPDTG 300
DB 241 GDTAERVLEMIQLLDLLEATDPCLEKFLGRIPMVENVVILTPHGYFAQDNVLYGYPDTG 300
QY 301 GQVVYILDOVRALENEMLRIRKQOGLNTPRLIITRLLPDAVGTTCQORLEKYGTSHS 360
DB 301 GQVVYILDOVRALENEMLRIRKQOGLNTPRLIITRLLPDAVGTTCQORLEKYGTSHS 360
QY 361 DILRVPRTEKGIYRKWISRFKWPYLETYTEDVAHEISKELHGTPLIIGNSDGNIV 420
DB 361 DILRVPRTEKGIYRKWISRFKWPYLETYTEDVAHEISKELHGTPLIIGNSDGNIV 420
QY 421 ASLAAHKLGVTOCTIAHALEKTKYPSDSIYWKLEEDKYHFSQOPTADLFAMNHTDFIITS 480
DB 421 ASLAAHKLGVTOCTIAHALEKTKYPSDSIYWKLEEDKYHFSQOPTADLFAMNHTDFIITS 480
QY 481 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVFPKFNIVSGADMEIYFPYTEKR 540
DB 481 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVFPKFNIVSGADMEIYFPYTEKR 540
QY 541 RLKHHPEIEDLLTYKYNEBEHLCVLNDRNKPIILFTMRDLRVKVLTLGLVEMCGKNPKLR 600
DB 541 RLKHHPEIEDLLTYKYNEBEHLCVLNDRNKPIILFTMRDLRVKVLTLGLVEMCGKNPKLR 600
QY 601 ELANLVVVGDDRRESKDLBEKAKMKQFELIDKYNLNGQFRWISSQMNRIIRNVELYRYI 660
DB 601 ELANLVVVGDDRRESKDLBEKAKMKQFELIDKYNLNGQFRWISSQMNRIIRNVELYRYI 660
QY 661 CDTKAFVQPALYEAFTLVVEAMTCGLPTPATCNGGPAEIIIVHKSQFNIDPHGDDAA 720
DB 661 CDTKAFVQPALYEAFTLVVEAMTCGLPTPATCNGGPAEIIIVHKSQFNIDPHGDDAA 720
QY 721 DILVDFPEKCKDPSHMDKISQGLKRIEKEYTWKISERLLTLTGYYGFPMKHSNLERR 780
DB 721 DILVDFPEKCKDPSHMDKISQGLKRIEKEYTWKISERLLTLTGYYGFPMKHSNLERR 780
QY 781 ESRRYLEMFYALKYRKLAESVPLAEE 806
DB 781 ESRRYLEMFYALKYRKLAESVPLAEE 806

RESULT 2
US-10-424-599-258404
; Sequence 258404, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 258404
; LENGTH: 805
; TYPE: PRT
; ORGANISM: glycine max
; APPLICANT: Screen, Steven E

OTHER INFORMATION: Clone ID: PAT_MRT3847_75363C.1.pep
US-10-424-599-258404
Query Match 88.6%; Score 3763.5; DB 12; Length 805;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 709; Conservative 49; Mismatches 47; Indels 1; Gaps 1;
QY 1 MAERALTIVSHLSRRELDLTAHNEILALISRIEGKGIQHQIILIEBAIPSENRK 60
DB 1 MATORLTRVSHLSRRELDLTAHNEILALISRIEGKGIQHQIILIEBAIPSENRK 60
QY 61 KLANGAFEEVLKASQGEAIVLPWPVALAVRPPGWMEYIRVNVHVLVVEELTVAEYLHPKE 120
DB 61 KLANGAFEEVLKASQGEAIVLPWPVALAVRPPGWMEYIRVNVHVLVVEELTVAEYLHPKE 120
QY 121 ELVNGSSNGNFVLELDEFPNSSFPRLTSGKSGVGFNLRHSAKLPHDKESHPLLE 180
DB 121 ELVNGSSNGNFVLELDEFPNSSFPRLTSGKSGVGFNLRHSAKLPHDKESHPLLE 180
QY 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEEYIGTLPPETPCAEFHRFOEIGLERGW 240
DB 181 FLRVCHGKGMMLNDRIQNLALOHVLRKAEEYIGTLPPETPCAEFHRFOEIGLERGW 240
QY 241 GDTAERVLEMIQLLDLLEATDPCLEKFLGRIPMVENVVILTPHGYFAQDNVLYGYPDTG 300
DB 241 GDTAERVLEMIQLLDLLEATDPCLEKFLGRIPMVENVVILTPHGYFAQDNVLYGYPDTG 300
QY 301 GQVVYILDOVRALENEMLRIRKQOGLNTPRLIITRLLPDAVGTTCQORLEKYGTSHS 360
DB 301 GQVVYILDOVRALENEMLRIRKQOGLNTPRLIITRLLPDAVGTTCQORLEKYGTSHS 360
QY 361 DILRVPRTEKGIYRKWISRFKWPYLETYTEDVAHEISKELHGTPLIIGNSDGNIV 420
DB 361 DILRVPRTEKGIYRKWISRFKWPYLETYTEDVAHEISKELHGTPLIIGNSDGNIV 420
QY 421 ASLAAHKLGVTOCTIAHALEKTKYPSDSIYWKLEEDKYHFSQOPTADLFAMNHTDFIITS 480
DB 421 ASLAAHKLGVTOCTIAHALEKTKYPSDSIYWKLEEDKYHFSQOPTADLFAMNHTDFIITS 480
QY 481 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVFPKFNIVSGADMEIYFPYTEKR 540
DB 481 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVFPKFNIVSGADMEIYFPYTEKR 540
QY 541 RLKHHPEIEDLLTYKYNEBEHLCVLNDRNKPIILFTMRDLRVKVLTLGLVEMCGKNPKLR 600
DB 541 RLKHHPEIEDLLTYKYNEBEHLCVLNDRNKPIILFTMRDLRVKVLTLGLVEMCGKNPKLR 600
QY 600 ELVNLVVVAGDDRRESKDLBEKAKMKQYGLIETKYNLNGQFRWISSQMNRIIRNVELYRYI 659
DB 600 ELVNLVVVAGDDRRESKDLBEKAKMKQYGLIETKYNLNGQFRWISSQMNRIIRNVELYRYI 659
QY 661 CDTKAFVQPALYEAFTLVVEAMTCGLPTPATCNGGPAEIIIVHKSQFNIDPHGDDAA 720
DB 661 CDTKAFVQPALYEAFTLVVEAMTCGLPTPATCNGGPAEIIIVHKSQFNIDPHGDDAA 720
QY 721 DILVDFPEKCKDPSHMDKISQGLKRIEKEYTWKISERLLTLTGYYGFPMKHSNLERR 780
DB 721 DILVDFPEKCKDPSHMDKISQGLKRIEKEYTWKISERLLTLTGYYGFPMKHSNLERR 780
QY 781 ESRRYLEMFYALKYRKLAESVPLAEE 806
DB 781 ESRRYLEMFYALKYRKLAESVPLAEE 806

RESULT 3
US-10-425-114-54297
; Sequence 54297, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E

APPLICANT: Tabaka, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 54297
LENGTH: 806
TYPE: PRT
ORGANISM: glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 700852943_Fli.pep
US-10-425-114-54297

Query Match 88.2%; Score 3744.5; DB 12; Length 806;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 705; Conservative 51; Mismatches 49; Indels 1; Gaps 1;

1 MAERALTIVHSRERLDELTLAHREITALLSRIGKGILOHIOIIEPAIEENRK 60
2 MATDRLTVHSRERLDELTLAHREITALLSRIGKGILOHIOIIEPAIEENRK 61
61 KLANGAFPEVLKASQEAIVLPWVALAVRPPGWVEYIRVNVHVALVVELTVAEYLHPE 120
62 KLTGAFPEVLKASQEAIVLPWVALAVRPPGWVEYIRVNVHVALVVELTVAEYLHPE 121
121 ELVDSSNGNFVLELDFEPNNSPPRPTLSKISGVGEFLNRHLSAKLPHDKESNAPLLE 180
122 ELVDSSNGNFVLELDFEPNNSPPRPTLSKISGVGEFLNRHLSAKLPHDKESNAPLLE 181
181 FLRVCHGKMMNDRIQNALOHVLRKAEYLGTPPTPCAFEFHROEIGLERGW 240
182 FLRLHSVKGKTMNDRIQNDALOHVLRKAEYLGTPPTPCAFEFHROEIGLERGW 241
241 GDTARVLEMIQLDLLEATDPTLEKFLGRIPWVFVNLITPHGYFAODNVLAGYPTG 300
242 GDMARVLESIOQLDLLEATDPTLEKFLGRIPWVFVNLITPHGYFAODNVLAGYPTG 301
301 GQVYVILDOVALLENEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLEKVGTEHS 360
302 GQVYVILDOVALLENEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLEKVGTEHS 361
361 DILRVPTTEGIVRKVTSRPEKWPYLETYTEVVAHEISKELHGTPLLIGNSDGNIV 420
362 HILRVPTTEGIVRKVTSRPEKWPYLETYTEVVAHEISKELHGTPLLIGNSDGNIV 420
421 ASLNAHKLGVQCTIAHLEKTKYPSDIIYMKLEDKYHSCOPTADLPANMHTDFIITS 480
421 ASLNAHKLGVQCTIAHLEKTKYPSDIIYMKLEDKYHSCOPTADLPANMHTDFIITS 480
481 TFOEIASGKDVGVYESHATAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEKR 540
481 TFOEIASGKDVGVYESHATAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEKR 540
541 RLKHFHPIEDLLATYKVENESHLVCLDRNKPIIFTMRLDRVKNITGLVEMYGNAKLR 600
541 RLKHFHPIEDLLATYKVENESHLVCLDRNKPIIFTMRLDRVKNITGLVEMYGNAKLR 600
601 ELANLVVVGSDRRKSKDLEKAEKMKFELIDKYNLNGOPRWISSQNNRIIRNVELYXI 660
601 ELANLVVVGSDRRKSKDLEKAEKMKFELIDKYNLNGOPRWISSQNNRIIRNVELYXI 660
661 CDTKGAFVOPALYEAFGTLVEAMTQGLPTFATCNGPAEIIIVHSGSFGNIDPYHGOAA 720
661 CDTKGAFVOPALYEAFGTLVEAMTQGLPTFATCNGPAEIIIVHSGSFGNIDPYHGOAA 720
721 DILVDFEKKCKDPSHMDKISGGGLIRBEKTKWIKYISERLITLTGVYGFMKVSNLERR 780
721 DILVDFEKKCKDPSHMDKISGGGLIRBEKTKWIKYISERLITLTGVYGFMKVSNLERR 780
781 ESRRLVEMFYALKYRKLAESVPLAE 806

Db 781 ESRRLVEMFYALKYRKLAESVPLAE 806

RESULT 4
US-10-424-599-154986
Sequence 154986, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 154986
LENGTH: 805
TYPE: PRT
ORGANISM: glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_110974C.1.pep
US-10-424-599-154986

Query Match 88.1%; Score 3743.5; DB 12; Length 805;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 705; Conservative 51; Mismatches 49; Indels 1; Gaps 1;

1 MAERALTIVHSRERLDELTLAHREITALLSRIGKGILOHIOIIEPAIEENRK 60
1 MATDRLTVHSRERLDELTLAHREITALLSRIGKGILOHIOIIEPAIEENRK 60
61 KLANGAFPEVLKASQEAIVLPWVALAVRPPGWVEYIRVNVHVALVVELTVAEYLHPE 120
61 KLTGAFPEVLKASQEAIVLPWVALAVRPPGWVEYIRVNVHVALVVELTVAEYLHPE 120
121 ELVDSSNGNFVLELDFEPNNSPPRPTLSKISGVGEFLNRHLSAKLPHDKESNAPLLE 180
121 ELVDSSNGNFVLELDFEPNNSPPRPTLSKISGVGEFLNRHLSAKLPHDKESNAPLLE 180
181 FLRVCHGKMMNDRIQNALOHVLRKAEYLGTPPTPCAFEFHROEIGLERGW 240
181 FLRLHSVKGKTMNDRIQNDALOHVLRKAEYLGTPPTPCAFEFHROEIGLERGW 240
241 GDTARVLEMIQLDLLEATDPTLEKFLGRIPWVFVNLITPHGYFAODNVLAGYPTG 300
241 GDMARVLESIOQLDLLEATDPTLEKFLGRIPWVFVNLITPHGYFAODNVLAGYPTG 300
301 GQVYVILDOVALLENEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLEKVGTEHS 360
301 GQVYVILDOVALLENEMLRIRKQGLNITPRILITRLLPDAVGTTCGRLEKVGTEHS 360
361 DILRVPTTEGIVRKVTSRPEKWPYLETYTEVVAHEISKELHGTPLLIGNSDGNIV 420
361 DILRVPTTEGIVRKVTSRPEKWPYLETYTEVVAHEISKELHGTPLLIGNSDGNIV 420
421 ASLNAHKLGVQCTIAHLEKTKYPSDIIYMKLEDKYHSCOPTADLPANMHTDFIITS 480
421 ASLNAHKLGVQCTIAHLEKTKYPSDIIYMKLEDKYHSCOPTADLPANMHTDFIITS 480
481 TFOEIASGKDVGVYESHATAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEKR 540
481 TFOEIASGKDVGVYESHATAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEKR 540
541 RLKHFHPIEDLLATYKVENESHLVCLDRNKPIIFTMRLDRVKNITGLVEMYGNAKLR 600
541 RLKHFHPIEDLLATYKVENESHLVCLDRNKPIIFTMRLDRVKNITGLVEMYGNAKLR 600
601 ELANLVVVGSDRRKSKDLEKAEKMKFELIDKYNLNGOPRWISSQNNRIIRNVELYXI 660
601 ELANLVVVGSDRRKSKDLEKAEKMKFELIDKYNLNGOPRWISSQNNRIIRNVELYXI 660

Db 600 ELVNLVAVGDRRKESKDLSEKAEKKYVGLIETVYKLNQGRWISSQMNRRVANGELRYVI 659
Qy 661 CDTKGAFVQPALYEAFGLTVEAMTCGLPTFATCNGGPAEIIYHGKSGFNIDPHGDOAA 720
Db 660 CDTKGAFAVQPALYEAFGLTVEAMTCGLPTFATCNGGPAEIIYHGKSGFNIDPHGDOAA 719
Qy 721 DILVDFEKKCKDPSHMDKISQGGKRIIEKYTWKIYSERLTLTGYYGFMKHSNLER 780
Db 720 DILVDFEKKCKDPSHMDKISQGGKRIIEKYTWKIYSERLTLTGYYGFMKHSNLER 779
Qy 781 ESRRYLEMFFALKYRKLAESVPLAE 806
Db 780 ESRRYLEMFFALKYRKLAESVPLAE 805

RESULT 5
US-10-424-599-154301
; Sequence 154301, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 154301
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110355C.1.pep
US-10-424-599-154301

Query Match 86.1%; Score 3656.5; DB 12; Length 806;
Best Local Similarity 85.2%; Pred. No. 0;
Matches 687; Conservative 59; Mismatches 59; Indels 1; Gaps 1;

Qy 1 MAERALTIVHSIRERLDETLTAHNEIILALSRIGKGGKILQHQIILEFPAIEENRK 60
Db 1 MANHPLTSHSFRERFDETLTGRNEIILALSRLEAKGKILQHQVAAEFEEIPEESRK 60
Qy 61 KLANGAFEEVYKASQEAIVLPWVALAVPRPGWERYRVNVHAIYVEELTVAEYLHPKE 120
Db 61 KLOGVFGEVLRSTQEAIVLPWVALAVPRPGWERYRVNVHAIYVEELTVAEYLHPKE 120
Qy 121 ELVDGSSNGNFVLELDFEFPNSSFPRLTSSKISNGVEFLNRHLSAKLFHDKESMHPLE 180
Db 121 ELVESSNGNFVLELDFEFPNSSFPRLTSSKISNGVEFLNRHLSAKLFHDKESMHPLE 180
Qy 181 FLRVHCHGKNNMLNDRIQNLALQHVLRKAEYVLTLPPTPCAEFHRFOEIGLERGW 240
Db 181 FLRLHSYKGTMMNDKQSLDSLOHVLRKAEYVLTVAPEPYSEFENKEREIGLERGW 240
Qy 241 GGTARVLEMIQLLDLEATDPCLEKLGRIIMVFNVVILTPHGYFAODNVLYGYPDG 300
Db 241 GGIARVLEMIQLLDLEATDPCLEKLGRIIMVFNVVILTPHGYFAODNVLYGYPDG 300
Qy 301 GGVVILLDQVRLAENEMLRIRKQGLINTPRILITRLLPDAVGTTCGQRLKRYGTHS 360
Db 301 GGVVILLDQVRLAENEMLRIRKQGLINTPRILITRLLPDAVGTTCGQRLKRYGTHS 360
Qy 361 DILRVFERTKGIYVRKMSRFEKWPYLETYEDVAHEISKEHGTPDLLITGNKSDGNIY 420
Db 361 DILRVFERTKGIYVRKMSRFEKWPYLETYEDVAHEISKEHGTPDLLITGNKSDGNIY 419
Qy 421 ASLLAHKLGVTQCTIAHALEKTYPDSDIYMKKLDKXHSQCFADLFAMNHTDFTITS 480
Db 420 ASLLAHKLGVTQCTIAHALEKTYPDSDIYMKKLDKXHSQCFADLFAMNHTDFTITS 479

Qy 481 TPQEIAGSKDTVGQYSEHTAFTLPGLYRVYHGDIVDPKKNITUSPGADMETYPYTEER 540
Db 480 TPQEIAGSKDTVGQYSEHTAFTLPGLYRVYHGDIVDPKKNITUSPGADMETYPYTEER 539
Qy 541 RLKHFHEIDILYTKYNESEHLCVLDNRNKPILFTMPRLDRVNLGLVWCKSNPKLR 600
Db 540 RLTFHFDIEBLYSVENEHEICVLDNRNKPILFTMARLDVRNITGLVEMYGKARLR 599
Qy 601 ELANLVVVGDRRKESKDLSEKAEKKYVGLIETVYKLNQGRWISSQMNRRVANGELRYVI 660
Db 600 ELVNLVAVGDRRKESKDLSEKAEKKYVGLIETVYKLNQGRWISSQMNRRVANGELRYVI 659
Qy 661 CDTKGAFVQPALYEAFGLTVEAMTCGLPTFATCNGGPAEIIYHGKSGFNIDPHGDOAA 720
Db 660 CDTKGAFAVQPALYEAFGLTVEAMTCGLPTFATCNGGPAEIIYHGKSGFNIDPHGDOAA 719
Qy 721 DILVDFEKKCKDPSHMDKISQGGKRIIEKYTWKIYSERLTLTGYYGFMKHSNLER 780
Db 720 DILVDFEKKCKDPSHMDKISQGGKRIIEKYTWKIYSERLTLTGYYGFMKHSNLER 779
Qy 781 ESRRYLEMFFALKYRKLAESVPLAE 806
Db 780 ESRRYLEMFFALKYRKLAESVPLAE 805

RESULT 6
US-10-425-114-50176
; Sequence 50176, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53113)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50176
; LENGTH: 811
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700852649_FLI.pep
US-10-425-114-50176

Query Match 86.1%; Score 3655.5; DB 12; Length 811;
Best Local Similarity 85.2%; Pred. No. 0;
Matches 687; Conservative 59; Mismatches 59; Indels 1; Gaps 1;

Qy 1 MAERALTIVHSIRERLDETLTAHNEIILALSRIGKGGKILQHQIILEFPAIEENRK 60
Db 6 MANHPLTSHSFRERFDETLTGRNEIILALSRLEAKGKILQHQVAAEFEEIPEESRK 65
Qy 61 KLANGAFEEVYKASQEAIVLPWVALAVPRPGWERYRVNVHAIYVEELTVAEYLHPKE 120
Db 61 KLOGVFGEVLRSTQEAIVLPWVALAVPRPGWERYRVNVHAIYVEELTVAEYLHPKE 125
Qy 121 ELVDGSSNGNFVLELDFEFPNSSFPRLTSSKISNGVEFLNRHLSAKLFHDKESMHPLE 180
Db 126 ELVESSNGNFVLELDFEFPNSSFPRLTSSKISNGVEFLNRHLSAKLFHDKESMHPLE 185
Qy 181 FLRVHCHGKNNMLNDRIQNLALQHVLRKAEYVLTLPPTPCAEFHRFOEIGLERGW 240
Db 186 FLRLHSYKGTMMNDKQSLDSLOHVLRKAEYVLTVAPEPYSEFENKEREIGLERGW 245
Qy 241 GGTARVLEMIQLLDLEATDPCLEKLGRIIMVFNVVILTPHGYFAODNVLYGYPDG 300
Db 241 GGIARVLEMIQLLDLEATDPCLEKLGRIIMVFNVVILTPHGYFAODNVLYGYPDG 300

```

Db      246 GDIAERVLEMIQLLDLLEADPCTLEFLGRVFWFNVVILSPHGYAODNVLGYPDTG 305
Qy      301 GOVVYILDOVRALLENEMLRIRKQOGLNTPRLITRLLPDAVGTTCORLEKYVGTSHS 360
Db      306 GOVVYILDOVRALLENEMLRIRKQOGLNTPRLITRLLPDAVGTTCORLEKYVGTSHS 365
Qy      361 DILRVPRTEKGIYAKWISREPEKWPVLETTYEDVAHEISKEHAGTPDLIIGNXSDGNIV 420
Db      366 DILRVPRTEKGIYAKWISREPEKWPVLETTYEDVAHEISKEHAGTPDLIIGNXSDGNIV 424
Qy      421 ASLNAHKIGVTOCTTAHALEKTKYPSDSIYWKKLEDEKTHFSQCFADLPFAMNHTDPIITS 480
Db      425 ASLNAHKIGVTOCTTAHALEKTKYPSDSIYWKKLEDEKTHFSQCFADLPFAMNHTDPIITS 484
Qy      481 TFOETAGSKDITVGOYESHATAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEER 540
Db      485 TFOETAGSKDITVGOYESHATAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEER 544
Qy      541 RLKHFHPEIEDLLYTKVENEHLCVLDNRNKPILFTMPRLDRVKNLTGLVENCGRNPKLR 600
Db      545 RLKHFHPEIEDLLYTKVENEHLCVLDNRNKPILFTMPRLDRVKNLTGLVENCGRNPKLR 604
Qy      601 ELANLVVVGDRRKSSKDLSEKAEKMKFELIDKTNLNGQFRMISQNMRIANVELYRYI 660
Db      605 ELANLVVVGDRRKSSKDLSEKAEKMKFELIDKTNLNGQFRMISQNMRIANVELYRYI 664
Qy      661 CDTKAFQOPALYEAFTLVVEAMTCGLPTFATCNGSPAEIIVHKSQGNIDPYGDQA 720
Db      665 CDTKAFQOPALYEAFTLVVEAMTCGLPTFATCNGSPAEIIVHKSQGNIDPYGDQA 724
Qy      721 DILVDFPEKCKDPSHMDKISQGLKRIEKKYTWKYSERLLTLGVYGFWMGVSNLEBR 780
Db      725 DILVDFPEKCKDPSHMDKISQGLKRIEKKYTWKYSERLLTLGVYGFWMGVSNLEBR 784
Qy      781 ESRRYLEMFYALKYRKLAESVPLAE 806
Db      785 ESRRYLEMFYALKYRKLAESVPLAE 810

```

RESULT 7

```

US-10-424-599-154300
; Sequence 154300, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 154300
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110354C.1.pep
US-10-424-599-154300

```

Query Match 86.0%; Score 3652.5; DB 12; Length 806;
 Best Local Similarity 85.4%; Pred. No. 0;
 Matches 688; Conservative 58; Mismatches 59; Indels 1; Gaps 1;

```

Qy      1 MAERLTVRHSLRERLDELTLAHRNEIALLSRIEGKGIGIHOIILEFAIPDENRK 60
Db      1 MANHVLTHSHSRERFDELTLGHRNEIALLSRLEKKGIGIHOIIVAFPEIIBESRK 60
Qy      61 KLANGAFPEEVLKASQEAIVLPPWVALAVPRPGWVEYLRVNVHALVEELTVAAVYLHPE 120
Db      61 KLQDGVGVGLASTQDAIVLPFVALAVPRPGWVEYLRVNVHMLVVDLRAEYLRFXE 120

```

```

Qy      121 ELVDSSNGNFVLELDFEPFNSSFPRLTSKISINGVEFLNRHLSAKLFHDKESMHPLLE 180
Db      121 ELVESSNGNFVLELDFEPFNASSFPRLTNSINGVEFLNRHLSAKLFHDKESQHPLLE 180
Qy      181 FLRVCHKQKMMMLNDRIQONALQHLVLRKAEVYGLTPETPCAEFEHRFOEIGLRGW 240
Db      181 FLRLHSYKQKTMMLNDKQSDLSLOHLVLRKAEVYLSVAPETPSEFEHRFREIGLRGW 240
Qy      241 GDTARVLEMIQLLDLLEADPCTLEFLGRVFWFNVVILSPHGYAODNVLGYPDTG 300
Db      241 GDTARVLEMIQLLDLLEADPCTLEFLGRVFWFNVVILSPHGYAODNVLGYPDTG 300
Qy      301 GOVVYILDOVRALLENEMLRIRKQOGLNTPRLITRLLPDAVGTTCORLEKYVGTSHS 360
Db      301 GOVVYILDOVRALLENEMLRIRKQOGLNTPRLITRLLPDAVGTTCORLEKYVGTSHS 360
Qy      361 DILRVPRTEKGIYAKWISREPEKWPVLETTYEDVAHEISKEHAGTPDLIIGNXSDGNIV 420
Db      361 DILRVPRTEKGIYAKWISREPEKWPVLETTYEDVAHEISKEHAGTPDLIIGNXSDGNIV 419
Qy      421 ASLNAHKIGVTOCTTAHALEKTKYPSDSIYWKKLEDEKTHFSQCFADLPFAMNHTDPIITS 480
Db      420 ASLNAHKIGVTOCTTAHALEKTKYPSDSIYWKKLEDEKTHFSQCFADLPFAMNHTDPIITS 479
Qy      481 TFOETAGSKDITVGOYESHATAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEER 540
Db      480 TFOETAGSKDITVGOYESHATAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEER 539
Qy      541 RLKHFHPEIEDLLYTKVENEHLCVLDNRNKPILFTMPRLDRVKNLTGLVENCGRNPKLR 600
Db      540 RLKHFHPEIEDLLYTKVENEHLCVLDNRNKPILFTMPRLDRVKNLTGLVENCGRNPKLR 599
Qy      601 ELANLVVVGDRRKSSKDLSEKAEKMKFELIDKTNLNGQFRMISQNMRIANVELYRYI 660
Db      600 ELANLVVVGDRRKSSKDLSEKAEKMKFELIDKTNLNGQFRMISQNMRIANVELYRYI 659
Qy      661 CDTKAFQOPALYEAFTLVVEAMTCGLPTFATCNGSPAEIIVHKSQGNIDPYGDQA 720
Db      660 CDTKAFQOPALYEAFTLVVEAMTCGLPTFATCNGSPAEIIVHKSQGNIDPYGDQA 719
Qy      721 DILVDFPEKCKDPSHMDKISQGLKRIEKKYTWKYSERLLTLGVYGFWMGVSNLEBR 780
Db      720 DILVDFPEKCKDPSHMDKISQGLKRIEKKYTWKYSERLLTLGVYGFWMGVSNLEBR 779
Qy      781 ESRRYLEMFYALKYRKLAESVPLAE 806
Db      780 ESRRYLEMFYALKYRKLAESVPLAE 805

```

RESULT 8

```

US-10-425-114-55184
; Sequence 55184, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack B
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55184
; LENGTH: 811
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700565776_F11.pep

```

US-10-425-114-55184

Query Match 86.0%; Score 3652.5; DB 12; Length 811;
Best Local Similarity 85.4%; Pred. No. 0;
Matches 688; Conservative 58; Mismatches 59; Indels 1; Gaps 1;

```
QY 1 MAERALTAVHSIRERLDETLTAHNEIALLSRIEGKGILOHHQIILEFEALPEENRK 60
DB 6 MANHPLTSHSFRERFDETLTGHRNEIALLSRLBAKGKGILOHHQVAAFEFEIPEESRK 65
QY 61 KLANGAFFEVUKAOGAIVLPBWVALAVRPPGWMEYIRUNVHALVVELVAEYLHKE 120
DB 66 KLQDVFGEVUKSTGEALVLPFVALAVRPPGWMEYIRUNVHALVVELVAEYLHKE 125
QY 121 ELVDSSNGNFVLEIDFEPFNSFPRLTSSKISGVGEFLNRHLSAKLFHDKESHAPLE 180
DB 126 ELVESSNGNFVLEIDFEPFNSFPRLTSSKISGVGEFLNRHLSAKLFHDKESHAPLE 185
QY 181 FLRVHCHGKMMMLNDRIQNLALQHVLRKAEEYLGTLPPETPCAEFEHFOEIGLERGW 240
DB 186 FLRLHSYKGTWMLNDKQSLDLSLOHVLARKAEEYLSVAPETPYSEFENRFRIGLERGW 245
QY 241 GDTARVLEMTOLLDLLEADPTCLEKFLGRIPVFNVVITLTPGYFAODNVLGYPDTG 300
DB 246 GDTARVLEMTOLLDLLEADPTCLEKFLGRIPVFNVVITLTPGYFAODNVLGYPDTG 305
QY 301 GOVVYIILDQVRLAENEMLRIRKQOGLNTPRILITRLLPDAVGTTCGRLKGVGTSHS 360
DB 306 GOVVYIILDQVRLAENEMLRIRKQOGLNTPRILITRLLPDAVGTTCGRLKGVGTSHS 365
QY 361 DILRVFPRTKGIYARKWISREKWPYLETYTEDEVAAHISKELHGTPLLIGNSDGNIV 420
DB 366 DILRVFPRTKGIYARKWISREKWPYLETYTEDEVAAHISKELHGTPLLIGNSDGNIV 424
QY 421 ASLNAHKIGVTOCTIAHALEKTKYPSDSIYMKKLEDKYHSCQFPADLPANMHTDFITTS 480
DB 425 ASLNAHKIGVTOCTIAHALEKTKYPSDSIYMKKLEDKYHSCQFPADLPANMHTDFITTS 484
QY 481 TFOEIASGNDTVGOYESHTAFTLPGLYRVVHIGIDVDFPKFNIVSPGADMEIYFPYTEER 540
DB 485 TFOEIASGNDTVGOYESHTAFTLPGLYRVVHIGIDVDFPKFNIVSPGADMEIYFPYTEER 544
QY 541 RLKHFHEIEDLLYTKVNEBEHLCVLNDNRKPIILFTMPRLDRVXNLTGLVEMCGNPKLR 600
DB 545 RLKHFHEIEDLLYTKVNEBEHLCVLNDNRKPIILFTMPRLDRVXNLTGLVEMCGNPKLR 604
QY 601 ELANLVVVGGRDRKSKDLBEKAEMKKMFELIDKYNLNGCFRWISSQNNRIRNVLYRYI 660
DB 605 ELVNLVVVAGDRRKSKDLBEKAEMKKMFELIDKYNLNGCFRWISSQNNRIRNVLYRYI 664
QY 661 CDTKGAFOPALYEAFGILVVEAMTCGLPTPATCNGSPAEIIVHKSQGFNIDPYHGDQA 720
DB 665 CDTKGAFOPALYEAFGILVVEAMTCGLPTPATCNGSPAEIIVHKSQGFNIDPYHGDQA 724
QY 721 DILVDFEKKCKDPSHMDKISQGLKRIIEKYTWKTIYSERLLTLTGYYGFWMKGVSNLERR 780
DB 725 ELVVEFEKSKADPSHMDKISQGLKRIIEKYTWKTIYSERLLTLTGYYGFWMKGVSNLERR 784
QY 781 ESRRLVEMFYALKYRKLAEVPLAE 806
DB 785 ESKRYLEMFYALKYRKLAEVPLAE 810
```

RESULT 9

US-10-137-036-77

; Sequence 77, Application US/10137036
; Publication No. US20030101478A1

; GENERAL INFORMATION:

; APPLICANT: Perera, Ranjan
; APPLICANT: Rice, Stephen
; APPLICANT: Bagleton, Claire
; APPLICANT: Lasham, Annette
; APPLICANT: Wood, Marion

```
; APPLICANT: Visser, Elizabeth  
; TITLE OF INVENTION: Compositions and Methods for the  
; FILE REFERENCE: 11000.1036c4  
; CURRENT APPLICATION NUMBER: US/10/137, 036  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: PCT/NZ 01/00115  
; PRIOR FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/724,624  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/598,401  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: PCT/NZ00/00018  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 60/146,591  
; PRIOR FILING DATE: 1999-07-30  
; PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/276,599  
; PRIOR FILING DATE: 1999-03-25  
; NUMBER OF SEQ ID NOS: 143  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 77  
; LENGTH: 805  
; TYPE: PRT  
; ORGANISM: Eucalyptus grandis  
US-10-137-036-77
```

Query Match 85.8%; Score 3643.5; DB 14; Length 805;
Best Local Similarity 85.0%; Pred. No. 0;
Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;

```
QY 1 MAERALTAVHSIRERLDETLTAHNEIALLSRIEGKGILOHHQIILEFEALPEENRK 60
DB 1 MADRLTSHSLRERLDETLTAHNRDIAFLSRVBAKGKGILOHHQIPAEFEALSEESRA 60
QY 61 KLANGAFFEVUKAOGAIVLPBWVALAVRPPGWMEYIRUNVHALVVELVAEYLHKE 120
DB 61 KLQDVFGEVUKSTGEALVSPWVALAVRPPGWMEYIRUNVHALVVELVAEYLHKE 120
QY 121 ELVDSSNGNFVLEIDFEPFNSFPRLTSSKISGVGEFLNRHLSAKLFHDKESHAPLE 180
DB 121 ELAOGSNGNFVLEIDFEPFNSFPRLTSSKISGVGEFLNRHLSAKLFHDKESHAPLE 180
QY 181 FLRVHCHGKMMMLNDRIQNLALQHVLRKAEEYLGTLPPETPCAEFEHFOEIGLERGW 240
DB 181 FLQVHCHGKMMMLNDRIQNLALQHVLRKAEEYLGTLPPETPCAEFEHFOEIGLERGW 240
QY 241 GDTARVLEMTOLLDLLEADPTCLEKFLGRIPVFNVVITLTPGYFAODNVLGYPDTG 300
DB 241 GDTARVLEMTOLLDLLEADPTCLEKFLGRIPVFNVVITLTPGYFAODNVLGYPDTG 300
QY 301 GOVVYIILDQVRLAENEMLRIRKQOGLNTPRILITRLLPDAVGTTCGRLKGVGTSHS 360
DB 301 GOVVYIILDQVRLAENEMLRIRKQOGLNTPRILITRLLPDAVGTTCGRLKGVGTSHS 360
QY 361 DILRVFPRTKGIYARKWISREKWPYLETYTEDEVAAHISKELHGTPLLIGNSDGNIV 420
DB 361 HILRVFPRTKGIYARKWISREKWPYLETYTEDEVAAHISKELHGTPLLIGNSDGNIV 419
QY 421 ASLNAHKIGVTOCTIAHALEKTKYPSDSIYMKKLEDKYHSCQFPADLPANMHTDFITTS 480
DB 420 ASLNAHKIGVTOCTIAHALEKTKYPSDSIYMKKLEDKYHSCQFPADLPANMHTDFITTS 479
QY 481 TFOEIASGNDTVGOYESHTAFTLPGLYRVVHIGIDVDFPKFNIVSPGADMEIYFPYTEER 540
DB 480 TFOEIASGNDTVGOYESHTAFTLPGLYRVVHIGIDVDFPKFNIVSPGADMEIYFPYTEER 539
QY 541 RLKHFHEIEDLLYTKVNEBEHLCVLNDNRKPIILFTMPRLDRVXNLTGLVEMCGNPKLR 600
DB 540 RLKHFHEIEDLLYTKVNEBEHLCVLNDNRKPIILFTMPRLDRVXNLTGLVEMCGNPKLR 599
QY 601 ELANLVVVGGRDRKSKDLBEKAEMKKMFELIDKYNLNGCFRWISSQNNRIRNVLYRYI 660
DB 600 ELANLVVVGGRDRKSKDLBEKAEMKKMFELIDKYNLNGCFRWISSQNNRIRNVLYRYI 659
```

```

Qy 661 CDTKGAFOAPALYEAFGTLTVEAMTCGLPTFATCNGCPAEIIVHGKSGFNIDPYHGDOA 720
Db 660 CDTKGVFOAPALYEAFGTLTVEAMTCGLPTFATCNGCPAEIIVHGKSGHIDPYHGDOA 719
Qy 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKIYSERLTLTGTVGFMKGVNLER 780
Db 720 ELVDFEKKCKIDQSHMDEISKGMQRIEKKYTWKIYSERLTLTAVVGFMKGVNLER 779
Qy 781 ESRRLMEFVALKYRKLAESVPLAE 806
Db 780 ESRRLMEFVALKYRKLAESVPLAE 805

RESULT 10
US-10-393-840-44
; Sequence 44, Application US/10393840
; Publication No. US20030229922A1
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE OF INVENTION: Modification of Plant Cell Wall Polysaccharides
; FILE REFERENCE: 11000.1012c3
; CURRENT FILING DATE: 2003-03-20
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 805
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-393-840-44

Query Match 85.8%; Score 3643.5; DB 15; Length 805;
Best Local Similarity 85.0%; Pred. No. 0;
Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;

Qy 1 MABRALTRVSHLRRLDETLAHRNEILALSRLEGKGLIQHQLLEFPAIDENRK 60
Db 1 MADMLTRSHSLRRLDETLAHRNDIVAFLSRVANKGKILQHQIFAEFEAISSESR 60
Qy 61 KLANGAFEEVVKASQEAIVLPWVALAVPRPGVWEYIRVNVHVLVVEELVAEYLHKE 120
Db 61 KLIDGAFGEVVKASQEAIVSPWVALAVPRPGVWEHVRNVHVLVLEQLVAEYLHKE 120
Qy 121 ELVDGSSNGNFVLELDEPPNSFPPTLSKISGVGEFLNRHSASLFDHKSHPLE 180
Db 121 ELADGSLNGNFVLELDEPPFASFPPTLSKISGVGEFLNRHSASLFDHKSHPLE 180
Qy 181 FLRVHCHGKMMMLNDRIQNLNALQHVLRKAEVYGLTPPTPCAEPFHRFOEIGLERGW 240
Db 181 FLQVHCHGKMMVMNARIQNVFSLQHVLRKAEVYLSLKPETPYSQFHFKEQOIGLERGW 240
Qy 241 GDTAEVLEMTQLLDLLEATDPTLKEFLGRIPVNFVNVLLTPHGYAONNVGYPTG 300
Db 241 GDTAEVLEMTQLLDLLEAPDPTLKEFLGRVNFVNVVWMSPHGYAODVVGYPPTG 300
Qy 301 GQVYVILDOVALENEMLRIKQGLNITPRILITRLLPDVAGTTCGRLEKYVTEHS 360
Db 301 GQVYVILDOVALENEMLRIKQGLDITPRILITRLLPDVAGTTCGRLEKYVTEHS 360
Qy 361 DILVDFEKKCKYRKMSRFEKWPVYETTEDVAHEISLHGTPLDILGNXSDGNTV 420
Db 361 DILVDFEKKCKYRKMSRFEKWPVYETTEDVAHEISLHGTPLDILGNXSDGNTV 420
Qy 420 HILRVFPRNEKGVVRKMSRFEKWPVYETTEDVAHEISLHGTPLDILGNXSDGNTV 419

```

```

Qy 421 ASILAHKLGVOCTTAHALEKTYPDSDIYWKLEKXYHFSQCFADLIFAMNHTDPIITS 480
Db 420 ASILAHKLGVOCTTAHALEKTYPDSDIYWKLEKXYHFSQCFADLIFAMNHTDPIITS 479
Qy 481 TFOEIAKSDTVQGVYESHMTAFTLPGLYRVHGVIDFDPKFNIVSPGADMEIYFPTBEKR 540
Db 480 TFOEIAKSDTVQGVYESHMTAFTLPGLYRVHGVIDFDPKFNIVSPGADMEIYFATBEOR 539
Qy 541 RLKHFPEIEDLLYTKVNEBHLCYLNRANKPILFTMRDLRVKVLTLGVKCGNPKLR 600
Db 540 RLKHFPEIEDLLYTKVNEBHLCYLNRANKPILFTMRDLRVKVLTLGVKCGNPKLR 599
Qy 601 ELANLVVVGGBRRKESKLEEKAKEMKPELIDKYNLNGOFWISSQNRIRNVLYRYI 660
Db 600 ELANLVVVGGBRRKESKLEEKAKEMKPELIDKYNLNGOFWISSQNRIRNVLYRYI 659
Qy 661 CDTKGAFOAPALYEAFGTLTVEAMTCGLPTFATCNGCPAEIIVHGKSGFNIDPYHGDOA 720
Db 660 CDTKGVFOAPALYEAFGTLTVEAMTCGLPTFATCNGCPAEIIVHGKSGHIDPYHGDOA 719
Qy 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKIYSERLTLTGTVGFMKGVNLER 780
Db 720 ELVDFEKKCKIDQSHMDEISKGMQRIEKKYTWKIYSERLTLTAVVGFMKGVNLER 779
Qy 781 ESRRLMEFVALKYRKLAESVPLAE 806
Db 780 ESRRLMEFVALKYRKLAESVPLAE 805

RESULT 11
US-10-393-840-144
; Sequence 144, Application US/10393840
; Publication No. US20030229922A1
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE OF INVENTION: Modification of Plant Cell Wall Polysaccharides
; FILE REFERENCE: 11000.1012c3
; CURRENT FILING DATE: 2003-03-20
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCT NZ/99/00169
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144
; LENGTH: 805
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-393-840-144

Query Match 85.8%; Score 3643.5; DB 15; Length 805;
Best Local Similarity 85.0%; Pred. No. 0;
Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;

Qy 1 MABRALTRVSHLRRLDETLAHRNEILALSRLEGKGLIQHQLLEFPAIDENRK 60
Db 1 MADMLTRSHSLRRLDETLAHRNDIVAFLSRVANKGKILQHQIFAEFEAISSESR 60
Qy 61 KLANGAFEEVVKASQEAIVLPWVALAVPRPGVWEYIRVNVHVLVVEELVAEYLHKE 120
Db 61 KLIDGAFGEVVKASQEAIVSPWVALAVPRPGVWEHVRNVHVLVLEQLVAEYLHKE 120
Qy 121 ELVDGSSNGNFVLELDEPPNSFPPTLSKISGVGEFLNRHSASLFDHKSHPLE 180
Db 121 ELADGSLNGNFVLELDEPPFASFPPTLSKISGVGEFLNRHSASLFDHKSHPLE 180
Qy 181 FLRVHCHGKMMMLNDRIQNLNALQHVLRKAEVYGLTPPTPCAEPFHRFOEIGLERGW 240
Db 181 FLQVHCHGKMMVMNARIQNVFSLQHVLRKAEVYLSLKPETPYSQFHFKEQOIGLERGW 240

```


Db 181 FLQVHCYKGMNMMNARIQNVFSLOHVRKAEEYLTSLKPEPYSOFEHKFOEIGLRGW 240
Qy 241 GDTAEVLEMTQLLDLEATDPCTLEKFLGRIPMVFNVLITPHGFAOONVLYGPTG 300
Db 241 GDTAEVLEMTQLLDLEATDPCTLEKFLGRIPMVFNVLITPHGFAOONVLYGPTG 300
Qy 301 GQVYVILDOVALENEMMLRIKQOGLNTPRILITRLLPDAVGTTCGRLKXVTEHS 360
Db 301 GQVYVILDOVALENEMMLRIKQOGLNTPRILITRLLPDAVGTTCGRLKXVTEHS 360
Qy 361 DILVFPFTEKGIYRKMSRPEKWPYLETTEVDVAHEISKELGTPLLIGNXSDGNIV 420
Db 361 HILVFPFTEKGIYRKMSRPEKWPYLETTEVDVAHEISKELGTPLLIGNXSDGNIV 419
Qy 421 ASLAAHKLGVTQCTAAHLEKTKYPSDSIYWKLEDKYHFCQFADLFAMNHTDFTTS 480
Db 420 ASLAAHKLGVTQCTAAHLEKTKYPSDSIYWKLEDKYHFCQFADLFAMNHTDFTTS 479
Qy 481 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVDFPKFNIIVSPGADMEIYFYTEER 540
Db 480 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVDFPKFNIIVSPGADMEIYFYTEER 539
Qy 541 RLKHPFPIEDLLYTKVNEBHLCVLDRNKPIILFTMRDLRVKNLTGLVWCGKNPKLR 600
Db 540 RLKHPFPIEDLLYTKVNEBHLCVLDRNKPIILFTMRDLRVKNLTGLVWCGKNPKLR 599
Qy 601 ELANLVVVGGRRKESKOLEEAKEMKMFELIDKYNLNGOFPMWISSOMNRIRNVELYRVC 660
Db 600 ELANLVVVGGRRKESKOLEEAKEMKMFELIDKYNLNGOFPMWISSOMNRIRNVELYRVC 659
Qy 661 CDTKGAFVOPALYEAFAGLTVVEAMTCGLPTFATCNGCPAEIIVHGKSGFNIDPYHGDQAA 720
Db 660 CDTKGAFVOPALYEAFAGLTVVEAMTCGLPTFATCNGCPAEIIVHGKSGFNIDPYHGDQAA 719
Qy 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKISERLITLTVGVGFMKVSNLERR 780
Db 720 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKISERLITLTVGVGFMKVSNLERR 779
Qy 781 ESRRYLEMFYALKYRKLAESVPLAEE 806
Db 780 ESRRYLEMFYALKYRKLAESVPLAEE 805

RESULT 12
US-10-425-114-53296
; Sequence 53296, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 53296
; LENGTH: 823
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700451030_Flt.pep
US-10-425-114-53296

Query Match 77.7%; Score 3299; DB 12; Length 823;
Best Local Similarity 76.9%; Pred. No. 7, 3e-303;
Matches 619; Conservative 83; Mismatches 101; Indels 2; Gaps 2;

Qy 3 ERALTVAHSIRERIDETLTAHNEILALLSRIECKGCIIOHNOIIEF-EATPEENRK 61
Db 14 DRLVSLRHSYRERIGSLASHAPNELVAFTRKLKLGMLQPHOILAYNNALIEABREK 73
Qy 62 LANGAFPEVKAQSEALVLPFWALAVRPRGWMEYRANVVAHVEELGTAEYLHPKEE 121
Db 74 LKQAFEDVLRQAQEAIVIPFWALAIRPRGWMEYRANVVAHVEELGTAEYLHPKEE 133
Qy 122 LVDSNGNENVLBELDFEPNFSPPRPTLSKISGVGFELNRHLSAKLPHDKESMPLLEF 181
Db 134 LVEEGPNNNFVLBELDFEPNFSPPRPTLSKISGVGFELNRHLSAKLPHDKESMPLLEF 193
Qy 182 LRVNCHGKMMNDRIQNALNLOHVRKAEEYLTSPREPCLAEFHRPOELGELGWC 241
Db 194 LRANHYGMNMMNDRIRSLALQALRKAEEHSTLQADTPYSEFHRPOELGELGWC 253
Qy 242 DTAERVLEMTQLLDLEATDPCTLEKFLGRIPMVFNVLITPHGFAOONVLYGPTG 301
Db 254 DTAERVLEMTQLLDLEATDPCTLEKFLGRIPMVFNVLITPHGFAOONVLYGPTG 313
Qy 302 QVYVILDOVALENEMMLRIKQOGLNTPRILITRLLPDAVGTTCGRLKXVTEHS 361
Db 314 QVYVILDOVALENEMMLRIKQOGLNTPRILITRLLPDAVGTTCGRLKXVTEHS 373
Qy 362 DILVFPFTEKGIYRKMSRPEKWPYLETTEVDVAHEISKELGTPLLIGNXSDGNIV 421
Db 374 DILVFPFTEKGIYRKMSRPEKWPYLETTEVDVAHEISKELGTPLLIGNXSDGNIV 432
Qy 422 SLAAHKLGVTQCTAAHLEKTKYPSDSIYWKLEDKYHFCQFADLFAMNHTDFTTS 481
Db 433 SLAAHKLGVTQCTAAHLEKTKYPSDSIYWKLEDKYHFCQFADLFAMNHTDFTTS 492
Qy 482 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVDFPKFNIIVSPGADMEIYFYTEER 541
Db 493 TFOEIASGKDTVGOYESHATAFTLPGLYRVVHGIDVDFPKFNIIVSPGADMEIYFYTEER 552
Qy 542 RLKHPFPIEDLLYTKVNEBHLCVLDRNKPIILFTMRDLRVKNLTGLVWCGKNPKLR 601
Db 553 RLKHPFPIEDLLYTKVNEBHLCVLDRNKPIILFTMRDLRVKNLTGLVWCGKNPKLR 612
Qy 602 ELANLVVVGGRRKESKOLEEAKEMKMFELIDKYNLNGOFPMWISSOMNRIRNVELYRVC 661
Db 613 ELANLVVVGGRRKESKOLEEAKEMKMFELIDKYNLNGOFPMWISSOMNRIRNVELYRVC 672
Qy 662 DTKGAFVOPALYEAFAGLTVVEAMTCGLPTFATCNGCPAEIIVHGKSGFNIDPYHGDQAA 721
Db 673 DTKGAFVOPALYEAFAGLTVVEAMTCGLPTFATCNGCPAEIIVHGKSGFNIDPYHGDQAA 732
Qy 722 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKISERLITLTVGVGFMKVSNLERR 781
Db 733 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKISERLITLTVGVGFMKVSNLERR 792
Qy 782 ESRRYLEMFYALKYRKLAESVPLAEE 806
Db 793 ESRRYLEMFYALKYRKLAESVPLAEE 817

RESULT 13
US-10-425-114-56696
; Sequence 56696, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO: 56696
 LENGTH: 823
 TYPE: PRT
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: Clone ID: LIB3069-052-F8_FLI.pep
 US-10-425-114-56696

Query Match 77.7% Score 3299; DB 12; Length 823;
 Best Local Similarity 76.9%; Pred. No. 7.3e-303;
 Matches 619; Conservative 83; Mismatches 101; Indels 2; Gaps 2;

3 BRALTRVSLRERDELTLAHRNEILALSRLEGKGILOHQTILFE-BAIPENRK 61
 14 DRVLSRLHSVRERIGDSLSAHPNELVAFTRLKNGKMLQPHQIAYENNAIPAEAREK 73
 62 LANGAFEEVLKASOBAIVLPPWVALAIVPRPGWVEYIRVNHALVVEELTVAEYIAFKEE 121
 74 LKDGAFEDVLRAGQBAIVLPPWVALAIVPRPGWVEYIRVNHALVVEELTVAEYIAFKEE 133
 122 LVDSGNGNFVLELDFEPPNSSFPRPTLSKISGVVEPLNRHLSAKLPHDKSMPLLEF 181
 134 LVEEGPNNFVLELDFEPPNSSFPRPTLSKISGVVEPLNRHLSAKLPHDKSMPLLEF 193
 182 LRVHCHKGNMMLNDRIQNLALQHVLRKAEYLTLPETPCAFEFERFOELGKGMG 241
 194 LRAHYKMTMTMLNDRIRSLALQALRKAEHLSLTLOADTPYSEFHHRFQELGKGMG 253
 242 DTAERVLEMIQLLDLEATDPTCEKFLGRIIPWVFNVLITPHGYFAQDNVLAGYPTGG 301
 254 DCAKRAQETIHLDDLLEAPDPSTLEKFLGITIPWVFNVLITPHGYFAQDNVLAGYPTGG 313
 302 QVVYILDOVRALLENMLRIKQGGINTPRILITRLLPDAVGTTCGRLKGVTEHSD 361
 314 QVVYILDOVRAMENMLRIKQGGINTPRILITRLLPDAVGTTCGRLKGVTEHSD 373
 362 ILRVFRTKEGIVRKWISRFKWPYLETYTEVDVAHEISKELHGTDLIIGXSDGNIVA 421
 374 ILRVFRTKEGIVRKWISRFKWPYLETYTEVDVAHEISKELHGTDLIIGXSDGNIVA 432
 422 SLAAHLGVTOCTIAHALEKTKYRPSDIYMKLEDKYHFSQPTADLFAMNHTDPIITST 481
 433 CLAAHKGVTHCTIAHALEKTKYRPSDIYMKLEDKYHFSQPTADLFAMNHTDPIITST 492
 482 FOEIAAGKOTVGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADSIYFPYTESHR 541
 493 FOEIAAGKOTVGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADSIYFPYTESHR 552
 542 LKHPPEIEDLLYTKVNEEHLCVLNDKRPILFTMPRLDRVKNLTGLVEMGKPKLRE 601
 553 LKHPPEIEDLLYTKVNEEHLCVLNDKRPILFTMPRLDRVKNLTGLVEMGKPKLRE 612
 602 LANLVVVGDRKESKDLLEKAEKMKMFELIDKYNLNGOFRWISSOMNRIRNVELYRITC 661
 613 LANLVVVGDRKESKDLLEKAEKMKMFELIDKYNLNGOFRWISSOMNRIRNVELYRITC 672
 662 DTGAFFVOPALYEAFTLVVEAMTCGLPTFATCNGGPAEIIIVHGSNGFNIDPYGDOAAD 721
 673 DTGAFFVOPALYEAFTLVVEAMTCGLPTFATCNGGPAEIIIVHGSNGFNIDPYGDOAAD 732
 722 ILVDFEKKCKDPSIMDKISQGLKRIEKTWKIYSERLLTLTGVYGFVKVSLYERRE 781
 733 ILVDFEKKCKDPSIMDKISQGLKRIEKTWKIYSERLLTLTGVYGFVKVSLYERRE 792
 782 SRRYLEMFYALKYRKLAEVPLAEE 806
 793 TRRYLEMLYALKYRKLAEVPLAEE 817

RESULT 14
 US-10-425-114-57139
 ; Sequence 57139, Application US/10425114

Publication No. US2004003488A1
 GENERAL INFORMATION:
 APPLICANT: Liu, Jindong
 APPLICANT: Zhou, Yihua
 APPLICANT: Kovalic, David K.
 APPLICANT: Screen, Steven E.
 APPLICANT: Tabaska, Jack E.
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(5313)B
 CURRENT APPLICATION NUMBER: US/10/425.114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO: 57139
 LENGTH: 823
 TYPE: PRT
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: Clone ID: 700221147_FLI.pep
 US-10-425-114-57139

Query Match 77.7% Score 3299; DB 12; Length 823;
 Best Local Similarity 76.9%; Pred. No. 7.3e-303;
 Matches 619; Conservative 83; Mismatches 101; Indels 2; Gaps 2;

3 BRALTRVSLRERDELTLAHRNEILALSRLEGKGILOHQTILFE-BAIPENRK 61
 14 DRVLSRLHSVRERIGDSLSAHPNELVAFTRLKNGKMLQPHQIAYENNAIPAEAREK 73
 62 LANGAFEEVLKASOBAIVLPPWVALAIVPRPGWVEYIRVNHALVVEELTVAEYIAFKEE 121
 74 LKDGAFEDVLRAGQBAIVLPPWVALAIVPRPGWVEYIRVNHALVVEELTVAEYIAFKEE 133
 122 LVDSGNGNFVLELDFEPPNSSFPRPTLSKISGVVEPLNRHLSAKLPHDKSMPLLEF 181
 134 LVEEGPNNFVLELDFEPPNSSFPRPTLSKISGVVEPLNRHLSAKLPHDKSMPLLEF 193
 182 LRVHCHKGNMMLNDRIQNLALQHVLRKAEYLTLPETPCAFEFERFOELGKGMG 241
 194 LRAHYKMTMTMLNDRIRSLALQALRKAEHLSLTLOADTPYSEFHHRFQELGKGMG 253
 242 DTAERVLEMIQLLDLEATDPTCEKFLGRIIPWVFNVLITPHGYFAQDNVLAGYPTGG 301
 254 DCAKRAQETIHLDDLLEAPDPSTLEKFLGITIPWVFNVLITPHGYFAQDNVLAGYPTGG 313
 302 QVVYILDOVRALLENMLRIKQGGINTPRILITRLLPDAVGTTCGRLKGVTEHSD 361
 314 QVVYILDOVRAMENMLRIKQGGINTPRILITRLLPDAVGTTCGRLKGVTEHSD 373
 362 ILRVFRTKEGIVRKWISRFKWPYLETYTEVDVAHEISKELHGTDLIIGXSDGNIVA 421
 374 ILRVFRTKEGIVRKWISRFKWPYLETYTEVDVAHEISKELHGTDLIIGXSDGNIVA 432
 422 SLAAHLGVTOCTIAHALEKTKYRPSDIYMKLEDKYHFSQPTADLFAMNHTDPIITST 481
 433 CLAAHKGVTHCTIAHALEKTKYRPSDIYMKLEDKYHFSQPTADLFAMNHTDPIITST 492
 482 FOEIAAGKOTVGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADSIYFPYTESHR 541
 493 FOEIAAGKOTVGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSPGADSIYFPYTESHR 552
 542 LKHPPEIEDLLYTKVNEEHLCVLNDKRPILFTMPRLDRVKNLTGLVEMGKPKLRE 601
 553 LKHPPEIEDLLYTKVNEEHLCVLNDKRPILFTMPRLDRVKNLTGLVEMGKPKLRE 612
 602 LANLVVVGDRKESKDLLEKAEKMKMFELIDKYNLNGOFRWISSOMNRIRNVELYRITC 661
 613 LANLVVVGDRKESKDLLEKAEKMKMFELIDKYNLNGOFRWISSOMNRIRNVELYRITC 672
 662 DTGAFFVOPALYEAFTLVVEAMTCGLPTFATCNGGPAEIIIVHGSNGFNIDPYGDOAAD 721
 673 DTGAFFVOPALYEAFTLVVEAMTCGLPTFATCNGGPAEIIIVHGSNGFNIDPYGDOAAD 732

Oy	722	IIVDFEFKCKDSDHMDKISOGGLKRIIEKYTKYISERLTLTGCVGFMGVGNLERRE	761
Dd	733	LIVDFEFCQCADPSHMHSKISOGGLQRIIEKYTKYISERLMTLTGCVGFMYGVNLERRE	792
Oy	782	SRRYLEMFYLLKYRKLAESYPLAE	806
Dd	793	TTRILEMFLYALKRYTMASTVPLAVE	817

RESULT 15
US-10-425-114-72813

; Sequence 72813, Application US/10425114
; Publication No. US20040034888A1
; Inventor: TROTT, TROY

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Steven E. Taback

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

1. TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21 (53313)B

FILE REFERENCE: 38-21(55313)B
CURRENT APPLICATION NUMBER: US/10/425.114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEO ID NOS: 73128

SEQ ID NO 72813

LENGTH: 823

TYPE: PRT

ORGANISM: Zea

FEATURE:

OTHER INFORMATION: Clone ID: LIB4172-051-C3_FLI.pep

US-10-425-114-72813

Query Match	77.7%;	Score 3299;	DB 12;	Length 823;
Best Local Similarity	76.9%;	Pred.No. 7.3e-303;		
Matches 619; Conservative	83;	Mismatches 101;	Indels 2;	Gaps 2;

OY	3	ERATLTHSRERIDEYLLAHREHIEALLSRLEGKGLIIONHILFE-BAIEENRK	61
Dd	14	DVLSRTHSRERIGDSLNHRELVAVTRKLKLOKMLONHIAIYVNNALIEAREK	73
OY	62	LANGAFEEVLKASOEALVLRPMVALAVRPRGVEXIRVNVHVLVEELVAYELHKEE	121
Dd	74	LKOGAFEDVLRRAOGEALVIRPMVALAIRPRGVEXIRVNVSELVEELVREYLOFEKQ	133
OY	122	LYDSSNGNVLELDEPPFNSSFRPTLTKSIGNGYERPLNRLSKLPHDSENHPLLEF	181
Dd	134	LVEGPGNNNVLELDEPPFNASPRPSTLSKISGNGQFNRRLSKLPHDESMYPLTNF	193
OY	182	LRVCHGKKMMMLNDRIOMNALONHLYRAAEYLGTLPEPCEAFERFOBIGLEROG	241
Dd	194	LRAHNYGMMMLNDRIKRSLSAQRLKREHILSTLOADTYSEFHNRFELGELKMG	253
OY	242	DTAERVLEMIQLLIDLEATDPCTLEKFLGRIMVENNVILTRPHGYFAODNVLGYPDTG	301
Dd	254	DKAKRAQETIHLIDLELLEADPSTLEKFLGTIPMVENVILSPHGYFAODNVLGYPDTG	313
OY	302	QVVYIILDOVALLENEMLLRIKQOGLNITPRILITRLLPDANGTTCOGRLERYGTEHS	361
Dd	314	QVVYIILDOVAMENEMLLRIKQOGLDITPRILITRLLPDANGTTCOGRLERYGTEHS	373
OY	362	ILRVFPRTEKGIYAKWISREPEKWPYLEYTEVVAHISEKLSHTPDILIGXSPDNIVA	421
Dd	374	ILRVFPRTEGIYAKWISRE-VMPILEYTTDVAHEINIGELOANPDLILIGXSPDNIVA	433
OY	422	SLLAHLKLVTOCTIAHALEKTKYPRDSDIYWKCKLEDKYHSCQPTADLPAHMHDTFIITST	481
Dd	433	CLLAHKMGVTHCTIAHALEKTKYPRNSDLVWKCFEDHYHSCQPTDILLAMHADIITST	493
OY	482	FOELASGKDYVGYEHSITATLPLGLRVVNGIDVPRPKNYIVSPGADMEIYFPTTEKGR	541

Dd	493	FOE1AGNDCTVGOQYESHNAFTWPGGLYRVVHGIDVDPPKFNIVSPGADLSITYEFTYESHKR	552
Qy	542	LKHREPREDELLTYVTEVENEHLCYLNDBDNKPLIFLTPMPRLDRVKNLTGLVEMWGSKNKLRE	601
Dd	553	LTSJLPELIEELLYSTENTTEHKEFVLNDRKPLIFSPARLDRKNKLTGLVGLGRNKLDE	612
Qy	602	LANLVVVGGDRRKESKDLEKAKEMKMFELLIDKYNLNGOPRWISSQONBIRNVELYRYIC	661
Dd	613	LVLNVVVGCGDHGNPSPCKDEQAEFFKMFELLIEQYNLNGHRIRWISSAQMNVRNDELRYIC	672
Qy	662	DTKGAFVOPALYEAGLTVVEMTQGLPFFATCNGSPRAIIVHSGSGFNIDRVHSGOADD	721
Dd	673	DTKGAFVOPALYEAGLTVVEMTQGLPFFATVAGGPABIIVHGSVGHNDIPVQGBDASA	732
Qy	722	ILVDFEKECKDPSHMDXISQGLKRIEBEYTWKTIYSELLTLTVGYVGFWKHVSNLERRE	781
Dd	733	LLVDFEFCQADPSHMSKISQCGLQRIEBEYTWKTIYSELLTMLTGVGYVFKVYSNLERRE	792
Qy	782	SRRYLMEFYALKYRCLASVPLAAEE	806
Dd	793	TKRYLEMLYALKYRTMASTVPLAAVE	817

Search completed: June 2, 2004, 14:54:51
Job time : 52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2004, 21:13:05 ; Search time 9925 Seconds
(without alignments)
7898.058 Million cell updates/sec

Title: US-10-003-405-1
Perfect score: 2625
Sequence: 1 atgagcagagcgtctccac.....tcaaaaaaaaaaaaaa 2625

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pln:*
20: em_ges_vit:*
21: em_ges_fun:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rod:*
26: em_ges_phg:*
27: em_ges_vrl:*
28: gb_ges1:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1302.4	49.6	3056 13	BU103683 SCCCR2100
2	1266.8	48.3	3140 11	AV103630 Zea mays
3	1085.2	41.3	2705 11	AV104856 Zea mays
4	696.6	26.5	1015 14	CD726832 EST028 Cu

5	677.8	25.8	689 13	BQ406651	BQ406651 GA_Ed009
6	677.2	25.8	689 12	BG439960	BG439960 GA_Ed000
7	674	25.7	683 13	BQ405173	BQ405173 GA_Ed007
8	664.8	25.3	746 14	CD486446	CD486446 CHR2_2F08
9	649.6	24.7	682 14	BQ406766	BQ406766 GA_Ed009
10	633	24.1	951 14	CF514627	CF514627 CABUd0005
11	630.2	24.0	656 13	BQ409552	BQ409552 GA_Ed002
12	623.2	23.7	637 13	BQ415686	BQ415686 GA_Ed010
13	623.2	23.7	664 9	AT731292	AT731292 BNGH1910
14	623	23.7	642 12	BG442235	BG442235 GA_Ed001
15	620.2	23.6	871 10	BE054845	BE054845 GA_Ed003
16	618.6	23.6	633 13	BQ415389	BQ415389 GA_Ed009
17	617.4	23.5	653 12	BG443072	BG443072 GA_Ed010
18	613.4	23.4	870 14	CD573686	CD573686 UCRPT01_0
19	610	23.2	634 14	BQ415266	BQ415266 GA_Ed009
20	608.6	23.2	672 10	AM561929	AM561929 IPRGH0001
21	596.2	22.7	945 14	CF213965	CF213965 CGF100081
22	581.8	22.2	881 14	CB893506	CB893506 EST646298
23	580.6	22.1	634 13	BQ407029	BQ407029 GA_Ed010
24	577.2	22.0	956 14	CK279779	CK279779 EST725857
25	573.8	21.9	585 13	BQ411883	BQ411883 GA_Ed004
26	573.6	21.9	947 14	CK274027	CK274027 EST720105
27	569	21.7	577 9	AT731115	AT731115 BNGH1873
28	568.4	21.7	928 14	CK284157	CK284157 EST746879
29	567.6	21.6	938 14	CK295709	CK295709 EST758423
30	567.2	21.6	688 12	BG445249	BG445249 GA_Ed002
31	560.6	21.4	921 14	CK285493	CK285493 EST748215
32	557	21.2	798 14	CD576633	CD576633 UCRPT01_0
33	555	21.1	685 13	BQ406489	BQ406489 GA_Ed009
34	553	21.1	999 10	BE416489	BE416489 MG0011.CO
35	550.6	21.0	933 14	CK279264	CK279264 EST725342
36	541.2	20.6	831 14	CF209749	CF209749 CAB20005
37	539.6	20.6	893 14	CK270584	CK270584 EST716662
38	536	20.4	693 14	CD486175	CD486175 CRRS7_1D0
39	535.4	20.4	799 14	CF118999	CF118999 MTU10CS.P
40	533	20.3	798 12	BG583977	BG583977 EST485737
41	530.4	20.2	797 14	CF118943	CF118943 MTU10CS.P
42	526.4	20.1	836 14	CF209223	CF209223 CAB20004
43	526	20.0	806 12	BG581937	BG581937 EST483673
44	525.4	20.0	806 14	CA917502	CA917502 EST641649
45	522.4	19.9	796 14	CB972247	CB972247 CAB10006_

ALIGNMENTS

RESULT 1
BU103683/LOCUS
DEFINITION SCCCR21002G07.g Saccharum officinarum mRNA (Nogueira,F.T.S)
ACCESSION BU103683
VERSION BU103683.1 GI:32815014
KEYWORDS EST.
SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE Nogueira,F.T.S., de Rosa,V.E. Jr., Menossi,M., Ulian,E.C. and Arruda,P.
AUTHORS RNA expression profiles and data mining of sugarcane response to low temperature
TITLE JOURNAL Plant Physiol. 132 (4), 1811-1824 (2003)
MEDLINE 22795309
PUBMED 12913139
COMMENT Contact: Nogueira FTS
Bioinformatics Lab
Organization for Nucleotide Sequencing and Analysis
C.P. 6176, Campinas, SP 13083-970, Brazil
Tel: 55 19 37881101
Fax: 55 19 37881089

Email: tebalid@unicamp.br.
Location/Qualifiers

1. .3056
/organism="Saccharum officinarum"
/mol_type="rRNA"
/db_xref="taxon:4547"
/clone_id="Saccharum officinarum mRNA (Nogueira, F.T.S.)"

ORIGIN

Query Match 49.6%; Score 1302.4; DB 13; Length 3056;
Best Local Similarity 71.3%; Pred. No. 2,2e-234;
Matches 1746; Conservative 0; Mismatches 697; Indels 6; Gaps 2;

```

QY 7 GAGCGTCTCACTCGGCTCCAGTCTCCGTGAGCGCTTGGATGAGACCTTCTTCT 66
DB GACCGTCTCGAGCGCGCTCCAGCGCTCAGGAGCGCATTTGGGATTCCTCTCC 2805
QY 67 CACAGGACGAGATTTGGCCCTGCTCCAGAGTCAGGGCCAAAGGAAAGAAATTCG 126
DB CACCCCAATGAGCTTGTGCGCTTTCACAGGCTGAAACCTTGGAAAGGATATGCTG 2745
QY 127 CAACACATCAATATTCTAGAGTTGA--AGCTATCCCTGAGAGAGAGAGAAAG 183
DB CAGCCCAACAGATATTGCGAGTACAAAGTCAATCCCTGAGGCTGAGCGTGAAG 2685
QY 184 CTCGTAATGTGCAATTTTGAAGTATGAAAGCTAGTCAGAAAGCAGTGTGCT 243
DB TTGAAGATGAGCGCTTGGAGATGCTGAGGCAAGCTCAGAGAAAGCAATTTATCCC 2625
QY 244 CCATGGTTCACCTGCTGCTGCTCCAGAGCGCTGCTTGGAGTACATTAAGTGAAT 303
DB CCATGGTTCACCTGCTGCTGCTCCAGAGCGCTGCTGCTGAGATATGTAAGGCTCA 2565
QY 304 GTTCAGCGCTTGTGTTGAGGAACTCATGTGCTGATCTCATCTTCAAGAGAG 363
DB GTTCAGAGCTGCTGCTGTTGAGAGCTGAGAGTCCAGAGTACCTGATTCAGAGAG 2505
QY 364 CTGTGTTGAGAACTTCAATGAACTTTGTTTGAATTGATTTGAGCCCTTCAAC 423
DB CTGTGTTGAGAAAGGCCCAACAACTTTGTTTGAAGTGAAGCTTGAAGCAATTCAT 2445
QY 424 TCATCATTCCTCCCGCCCACTTTCAAAATTCATGTTATGATGATGATGATGAT 483
DB GCATCTTCTCCCGCTTCTCTGTCAAAGTCAATGTCAAATGATGATGATGATGATG 2385
QY 484 GGTCACTTTGCGCAAAATTTGTCATGACAAAGAGACATGACCTTTGCTCGAATTC 543
DB AGGCACTGTATCAAGCTCTTCAATGACAAAGAGACATGATGATGATGATGATGATG 2325
QY 544 CTCAGAGTCAATTTGTCAGAAAGGCAAAATGATGTTGAATGACAAATTCAGAACTTG 603
DB CTTCGCGCCCACTCAAGAGGAGTACATGATGTTGAATGACAAATTCGCGAGTCTC 2265
QY 604 AATGCTTTTCAAGATTTTGAAGAAAGCAGAGAGTATCTTGTGATACCTTCTGAG 663
DB AGTGTCTGCAAGGCGCTCTGAGAAAGGCTGAGAGCACTGTTCACCTTCCAGCTAT 2205
QY 664 ACACATGTGCGAATTTGAAACACCGGTTCCAGAAATCGGTTTGAAGAGTTGGGT 723
DB ACCCATACTGTGAATTTCAACAGGTTCCAGAACTTGTGCTGAGAAAGGTTGGGT 2145
QY 724 GACACCGGAGAAAGCGTCTCGAGATGATCAATCTCTTTTGAATCTTCTTGAAGCACT 783
DB GACTGCGCTAAGCGCGCCAGAGAGTATTCACCTCTTTCGACCTTTCAGAGCCCA 2085
QY 784 GATCTTTCGACCTTGAAGAGTCTTGGGAGAAATCCCAATGATGATGATGATGAT 843
DB GATCTTTCGACCTTGAAGAGTCTTGGGAGATCCCATGATGATGATGATGATGATG 2025
QY 844 CTCATCTCCCAAGATATCTGCTCAAGACAAATGTTTGGGATATCCGACACCGTGG 903
DB CTCATCTCCCAAGATATCTGCTCAAGACAAATGTTTGGGATATCTGCTCAAGAC 1965
QY 2024 CTCCTCCCTCATGTTTACTTGTCTCAAGCTAATGCTTGGGTACCTTGAACCTGAGG 1965

```

```

QY 904 CAGGTTGTTATCATCTTTCGATCAAGTCCGAGCTTTGGAGATGATGATGATGATGAT 963
DB CAGGTTGTTATCATCTTTCGATCAAGTCCGAGCTTTGGAGATGATGATGATGATGATG 1964
QY 964 AAGCAACAGGATCAATCATCAACCCCTGAAATCTCTATATATATATATATATATATAT 1023
DB AAGCAATGTGTCTTTCGATCAATCAACCAAGATCTTATTTGTTACAGGTTCTCCCTGAT 1845
QY 1024 GCTGTGGAACCAATCATGCGCTCAACGATGAGAAAGTATACGGAACAGACACTCGGAT 1083
DB GCAATGGAACCAATCTGTGTGATGAGAGGCTTGAAGAGGCTTGGACATGACATGCGCAT 1785
QY 1084 ATTTTTCGATACCTTTCAGAACGAAAGGAAATTTGTTGAAATGATCTCAAGATTT 1143
DB ATCTTTCGATGATCTTTCAGAACGAAAGGAAATTTGTTGCAAGTATCTCAAGTCTT 1725
QY 1144 GAAAGTGTGCGCATCTTTCGAAACCTGACAGAGATGTTGTCATGAATCTCCAA 1203
DB G--AAAGTGTGCGCATCTTTCGAAACCTGACAGAGATGTTGTCATGAATCTCCAA 1724
QY 1204 GAGTTGACGCGCACGCGATCTGATCATCGGAAACNACAGCGACGCGATATGCTGCC 1263
DB GAGCTTACGCGCATCTGATCTGATCATGAGAACTTACAGTGAAGGAACTTGTGCG 1608
QY 1264 TCTTCTGCGACATTAATTAAGTGTCAACAGTGCACATCGCCCATGCTTTGAGAG 1323
DB TGTTCCTGCGCACATTAATTAAGTGTCAACAGTGCACATCGCCCATGCTTTGAGAG 1548
QY 1324 ACAAAATATCCAGATTCAGATATCTATTTGAAAGAGCTGAAAGCAATATCATTTCT 1383
DB ACTAATATCCAGATTCAGATATCTATTTGAAAGAGCTGAAAGCAATATCATTTCT 1547
QY 1384 TGCCAAATTCAGCTGATCTTTTTCAGATGAAACATACAGATTTATCATATCAAGTACT 1443
DB TGCCAAATTCAGCTGATCTTTTTCAGATGAAACATATGATGATGATGATGATGATG 1487
QY 1444 TTCCAGAAATTTGAGAGAGAGAGACATGTTGTCATATGAGAGGACATGCTTTC 1503
DB TTCCAGAAATTTGAGAGAGAGAGACATGTTGTCATATGATGATGATGATGATGATG 1427
QY 1504 ACTTTCCTGCTCTCAACCGGTTGATCATGATGATGATGATGATGATGATGATGATGATG 1563
DB ACTTTCCTGCTCTCAACCGGTTGATCATGATGATGATGATGATGATGATGATGATGATG 1367
QY 1564 ATTGTTTCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1623
DB ATTGTTTCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1307
QY 1624 TTGAAGATTTCCATCTGAGATGAGAACTTTTTCACCAAGATTTGAGATGAGAA 1683
DB TTGAAGATTTCCATCTGAGATGAGAACTTTTTCACCAAGATTTGAGATGAGAA 1247
QY 1684 CACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1743
DB CACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1187
QY 1744 GGTGTCAAGAACTTAAACCGGATCTGTCAGAGTGTGCGGCAAGAACCAAGTTGCTGAG 1803
DB GGTGTCAAGAACTTAAACCGGATCTGTCAGAGTGTGCGGCAAGAACCAAGTTGCTGAG 1127
QY 1804 TTGCTAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1863
DB TTGCTAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1067
QY 1864 AAGCTGAAATGAGAGAAATTTGAGCTGATGACAGTATCACTTTCAGAGAAAGAGAG 1923
DB AAGCTGAAATGAGAGAAATTTGAGCTGATGACAGTATCACTTTCAGAGAAAGAGAG 948
QY 1924 AGATGATATCATCTCAATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1983
DB AGATGATATCTCAATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 947
QY 1984 GACAGAAAGGTGCTTTTGTACAGCTGATGATGATGAGATGATGAGATGATGAGATG 2043

```


Oy	961	TTAAAGCAACAAGAGCTCAACATCAACCCCTGGAATCCCTATATATCAAGACTTCCT	1020
Db	1138	ATCAGACAGTGTGCTTGACATCACGCCGAAGATCTTATTGTGACCAAGTGTCTCCT	1197
Oy	1021	GATGCTGTGGAAACAACATG-CGGTCAACGACTTGAGAAAGTATACGGAACAGACCTC	1079
Db	1198	GATGCAACTGGACACCACTGTGTGGCCAGCGCTTGAGAAAGTCTTGGCACCGAGCACTG	1257
Oy	1080	GGATTTCTTGGAGTACCTCTCAGAACAGAAAAGGAAATTGTTCGAAAATGGAATCTCAAG	1139
Db	1258	CCATATCTTGGCGCGGCATTCAGAAACGAAAACGAATCGTTCGCAAGTGAATCTCCG	1317
Oy	1140	ATTGGAATAAGTCTGGCCATCACTTGGAAACCTACACAGAGGATGTGCTCATGAATCTC	1199
Db	1318	ATTG--AAGTCTGGCCGTACTGGAACCTTACCTAATGACGTGGCGCATGAGATGTC	1374
Oy	1200	CAAAAGTTGACACGGCACCGCAGATCTGATCATCGGAAACNACAGCGCAATATCTG	1259
Db	1375	TGAGAGCTTACAGGCAATCTCGACTGATCATGGAAAACATACGTGACGGAACTTGT	1434
Oy	1260	CGCTCTTGTCTCGCACATAAATTAGTGTCAACAGTGCACCATCGCCATGCTTTGGA	1319
Db	1435	TGCGGTGTGCTCGCCCAAGATGGGTGTACTCACTGATCAATGCGCCATGCGCTTGA	1494
Oy	1320	GAAAGCAAAATATCCAGATTGAGATATGTAATTGAAAGAAAGCTGGAAGCAATTCATTT	1379
Db	1495	GAAAACTAAGTACCTTAACCTCCGACTTACTGGAAGAAAGTTGAGATCACTACACTT	1554
Oy	1380	CTCTTGCAATTTACAGCTGATCTTTTGGCAATGACATACATACAGATTTCAATCAACAG	1439
Db	1555	CTCGTGCAAGTTCAACCACTGACTGATTGCAATGAACATCGCACTTCAATCATCACAG	1614
Oy	1440	TACTTTCAGAAATTGACAGAGCAAGACACTGTTGTCAATACGAGACCACTGC	1499
Db	1615	TACCTTCCAAAGATCGCCGGAACAAAGAACACCGTCGGCCAGTACGAGTCAACATGGC	1674
Oy	1500	TTTCACTTTCTCGTCTCTCAACCGGTGTGTACATGATATCGATGTTGTGATCCCAATT	1559
Db	1675	GTTCAACATGCTGTGCGCTGTACCCGGTGTCCACGGCAATTGATGTGTGACCCCAATT	1734
Oy	1560	CAACATTTTCCCCCTGTGCTGATATGAGATATTACTTCCCTTACACGGAAGAAACG	1619
Db	1735	CACATCTGTCTCTTGGCGGGACCTGTCACTTCTTCCGTACACCGAGTGGCAAA	1794
Oy	1620	GAGTTGAAGCAATTCATCTCTGATCGAAGACCTTTTACACAAAGTTGAGAATGA	1679
Db	1795	GAGGCTGACCTCCCTTCAACCGGAATGAGAGCTCTGTACAGCCAAACGAGAACAC	1854
Oy	1680	AGAACACTTATGTGTCTCAATGACCGCAACAGCATTTCTGTTCACAAATGCCAAGCT	1739
Db	1855	GGAGCACAAAGTCTGTCTGAACGACAGAAACAAAGCAATCATCTTCTCAATGCTCGCT	1914
Oy	1740	TGATGTATCAAGAACTTAACCGGACTGTGAGTGTGATGAGCGAAAGAAATCTAAAGATTGGA	1799
Db	1915	CGACCGTGTGAAGACTTGAACCTGAGCTGTGAGCTGTACCGCCGGAACAAAGCGCTGCA	1974
Oy	1800	TGAGTTGCTAACCTCGTATGTTGATGATGATGAGCGAAAGAAATCTAAAGATTGGA	1859
Db	1975	GGAGCTGTGAACCTGTGTGTGTGTGCGGGGACCATGGAACCTTCAAGAACAAAGGA	2034
Oy	1860	AGAGAGGCTGAATGAAGAAATTTGAGTGTATGCAAGTACCACTTGAACGCCCA	1919
Db	2035	GGAGCAGGCCGAGTTCAAGAAAGTGTTGACCTCATCAGACAGTACCACTGGAACGGGCA	2094
Oy	1920	ATTGAGATGGAATCATCTCAATGAATGAATCCGAATGTGAACCTTACCATCAT	1979
Db	2095	CATCCGCTGAGATCTCGCCAGATTAACCGGCTCGCAACGGCGAGCTGTACCCGTACAT	2154
Oy	1980	TTGGCACAAGAAAGGTGCTTTGTATCAGACTTGCACTGTATGAAAGCTTTGATTTGACAGT	2039
Db	2155	CTGGCACAACAAAGGCGCTTCTGTGTACGCTGCTTTTCTACAGAGCTTTGGGGCTGACGCT	2214
Oy	2040	TGTGAGGCAATGACTTGTGGTTGGCAACATTGGCAACTGTGAACGTGACAGGCCA	2099

Db	2215	GGTTGAGCCATGACCTGCGGCGCTGGCCACGTTTGGACAGACCTTACGGCGGTCGGCGGA	2274
Qy	2100	GATTATTGTGCATGAGGAATCTGGTTTCAACATTGATTCCTTACCATGTGTATCAAGCTGC	2159
Db	2275	GATCATGTGCACGCGCGTGTCTGGCTACACATCGACCTTACCGAGGCGCAAGGCGTC	2334
Qy	2160	TGACATACCTGTCGATTTCTTTGAAAAGTGAAGAAAGATTCATCTCACTGGGATTAAGAT	2219
Db	2335	GGGCTCTCTGTGGACTTCTTCGACNAGTCCAGGCGGACCCGAGCCACTGAGCAAGAT	2394
Qy	2220	CTCCCAAGAGAGCTTGAACGAATAGAGAGAAATATACATGGAAGATTACTTGGAGAG	2279
Db	2395	CTCCCAAGGAGGCGGTCCAGCGTATCGAGGAAGTACACCTGGAAGCTTACTCGAGAG	2454
Qy	2280	ACTATTGACCTTGACAGAGATGTATGGATTCTGGAGCATGTTTCCAACCTTGAACCGC	2339
Db	2455	GCTGATGACCTTCAACCGCGCTGTACGCGGTTCTGGAAGTACGTGTCCAACTGGAGAGCGC	2514
Qy	2340	TGAGAGTCGTGCTTACTCTTGAGATGTTTATGCTTAACTAAGTACCGTAAGCTGGCTAATC	2399
Db	2515	CGAACCCTGGGCGGTACCTGAGATGCTGTACGCGCTCAAGTACCGCACCATGGCGAGAC	2574
Qy	2400	AGTTCATTGGCAGAGAG 2418	
Db	2575	CGTCCCGCTGGCGGTGAG 2593	
RESULT 3			
LOCUS	AY104856	2705 bp	mRNA linear HTC 16-OCT-2002
DEFINITION	Zea mays PC0079420 mRNA sequence.		
ACCESSION	AY104856		
VERSION	AY104856.1	GI:21207934	
KEYWORDS	HTC.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Bukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 2705) Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.		
TITLE	Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes		
JOURNAL	Unpublished (2002)		
REFERENCE	2 (bases 1 to 2705)		
AUTHORS	Coe,B.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDb and may be found by BLAST searching at MSL, maizemap.org ; ZmDb, www.zmdb.iastate.edu ; TIGR, www.tigr.org ; or NCBI, www.ncbi.nlm.nih.gov . When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDb: www.zmdb.iastate.edu .		
FEATURES	location/Qualifiers		
SOURCE	1..2705		
	/organism="Zea mays"		
	/mol_type="mRNA"		
	/db_xref="MaizEDB:634420"		
	/db_xref="taxon:4577"		
	/clone_lib="Maize Mapping Project/Dupont Consensus Library"		
	/note="this sequence is part of a project of EST assemblies resulting from the application of public config to seed Dupont configs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"		

Query Match 41.3%; Score 1085.2; DB 11; Length 2705;
 Best Local Similarity 66.5%; Pred. No. 1,2e-193;
 Matches 1587; Conservative 0; Mismatches 794; Indels 6; Gaps 2;

38 GTGAGCGTTTGGATGAGACCTTCTTGTCTCAGAGACGAGATTTTGGCTGTCTCA 97
 1 GCGACCGGGTGGAGAGACCTTCCACGGGACCGGACGAGCTGTGGCCCTCTGTCA 60
 98 GGATGAGGGCAAGAGAGAAATTTGCAACACCATCAATTTTCTAGATTGAG 157
 61 AGACGTAAACAGGGGAGGAGCATCTGCAAGCCGACACATCTTCAGCGGCTCAG 120
 158 CTATCCCTGAGAGAGACAGAAAGAGCTGCTAATAGTGCAATTTTGAAGATTGAG 217
 121 AGGTCAAGGGCTCCGGGGGGCGCGCTAGCCGAGAGACCTTCTCCAGCTCTCCGCT 180
 218 CTAGTCAGAGAGAGCATCTGTGTGCTTCAATGGGTTGCACTTGTCTCCAGGCGCTG 277
 181 CCGGCAAGAGGCGATCGTGTGCGCGCTGTGTGGCCATGGCGGTGGCGCGCGCGG 240
 278 GTGTTGGAGATCAATGAGTGAATGTTACGCGCTTGTGTGAGAACTCACTGTG 337
 241 GAGTTGGAGATCACTCGCGCTCAACGTTCAAGAGCTCAGCTCAGAGCTCACAGTCT 300
 338 CTGAGTACTCCACTTCAAGAGAGAGCTTGTGATGAGAACTTCAATGAGAACTTGT 397
 301 CGAGTACCTCCGCTTCAAGAGAGAGCTTGTGAGCGGACAGACATGATCCCTACGTT 360
 398 TGAATGGAATTTTGGAGCTTCACTCATCTATCCCGGCGCACTCTTCAAAATCA 457
 361 TCGAGCTTGAATTCAGAGCTTCAATGTCTGAGTCCAGGCCAATCGGTATATCTA 420
 458 TTGTAATGATGATGAGTTCTTAAATCTGACCTTTCGCAAAATTTGTTCAATGAG 517
 421 TTGGAAGCGTGTGAGTTCTTCAACCGACACTTGTCTCAATGATTTCCGCAACAGG 480
 518 AGAGATGACACCTTGTCTCGAATCTTCAGAGTCAATTTGTCAAGAGGAGAGACGA 577
 481 ATGCTTGAAGCGCTGTGATGATTTCTCTGCGCACCGGACAGAGGCGATGTTATGA 540
 578 TGTGAATGACAGAAATTCAGAACTTGAATGCTTCAACATGTTTGAAGAAAGAGAG 637
 541 TCTTAAATGATGATTAAGAACTTGGGAGGCTTCACTGTGTGACCAAAAGCTGAG 600
 638 AGATATCTGTGACCTTACTCTCTGAGACACATGTGCGGAATTCAGAACCGGTTCA 697
 601 AGCACTTGTCAAAAGCTCCCTGTCTGACACACATCTCACAATTTGCTTAATAATTC 660
 698 AAATGCTTTGGAAGAGTTGGGGTGAACCGGAGAGAGCGTGTCCAGATATCCAA 757
 661 AGTGGGGCTGAGAGAGTTGGGGTGAATGAGAGAGGAGATTTTGAAGAAATTCATC 720
 758 TCTTTTGGATTTCTTGGAGCAATGATCTTGGACCTTGAAGTTCTTGGAGAA 817
 721 TCTTTTGAACATCAATTCAGGCGCAAGCCATCTACATGAGAAATTTCTGGGAGAG 780
 818 TCCCATGCTTCAATGATTTGATTTCTCACTCCGAGAGATCTTGTCTCAAGCAATG 877
 781 TCCCATGATTTTAAAGTTGTTGATTTCTCCCTCATGATATCTTGGTCAAGCTAATG 840
 878 TTTTGGGGTATCCGACACCGGTGGGAGTTTGAATCTTGAATCAAGTCCGAGCTT 937
 841 TATTAGGCTTGCACACAGAGAGAGAGATGCTTAATTAATGAGCAAGTTCTGTGAC 900
 938 TGAAGATGAGATGCTCTCCGTATTAAGCAAGAGATCAACATCAACCTCCGATCC 997
 901 TAGAAATGAGATGATGCTTCTTAAAGAAACAGAGGCTTGAATTTCCCAAGATTC 960
 998 TCAATTAATTAAGACTTCTCTGATGCTGTGAGAAACAATGCGGTCAACGATTTGAG 1057
 961 TCAATTTACTCGGCTGATACAGATGCAAAAGAGAAATCATGATCAGCGGCTTGAGA 1020

1058 AAGTATACGAGACAGAGCATCTCGATATTTCTGAGTACCTTCAGAACAGAAAGAGAA 1117
 1021 GAATTAAGAGAAACAGAGATACATTAATTAACAGGTTCCCTTGAGAAATGAAATGCGA 1080
 1118 TTGTTGAAATGATCTCAAGATTTGAAAAAGTCTGGCCATCTTGGAAACCTACACAG 1177
 1081 TACTTAAGAAATGATATCAAGATTTG---ATGTTGCGCAATCTGGAACATTTGTGTG 1137
 1178 AGGATGTTGCTCATGAAATCTCAAGAGTTGCAACGCGACCGCATGTGATCATCGGA 1237
 1138 AGGATGCTGCTGTAATTTCTGTGTAATTAACAGGATCTCAGACTTCAATATTGAA 1197
 1238 ACNACAGCGAGCGCAATATCTGCTCTCTTGTCTGCAATATAATTAAGTGTACACAGT 1297
 1198 ACTACAGGATGAGAAATTTGTGGGTATGCTATCTTAAGATGGAATTAACCGAGT 1257
 1298 GCACCATGCGCCATGCTTGTGAGAGAGACAAATATTCAGATTCAATATCTATTGAGAA 1357
 1258 GCAACATTTGCTCATGCTGTAGAAAAAGATAAGTATCCAGATTCAAGCATATTTGGA 1317
 1358 AGCTTGAAGAGAAATACATTTCTTGTGCAATTTAGAGCTGATCTTTTGCATAGAAC 1417
 1318 ATTCGATGAGAGAGATCAATTTCTCTGTCACTTCACTGATATTAATGCTAGAAACA 1377
 1418 ATACAGATTTCAATCAACAGATCTTTCAGAAATTCAGAGAAAGAGACAGACCTGTTG 1477
 1378 ATGCTGATTTTATCATCAACAGACATACAGAAATGCTGGAAGAGAAAATCTGTTG 1437
 1478 GTCAATACGAGAGCCACATCTCTTCACTTCTGTGCTCTACCGTGTGTAATGATA 1537
 1438 GACAGTATGAGAGATCACTGCTTTTACTGCTGCTGTGATCAGAGATTTGCAATGAGA 1497
 1538 TCGATGCTTGTGATTCGCAAAATTTCAATTTTCCCGGTGCTGATATGAGATATAC 1597
 1498 TCGATGCTTGTGATTCGCAAAATTTCAATTTTCCCGGTGCTGATATGAGATATAC 1557
 1598 TCCCTTACACGAGAGAGAGCGGAGTTGAAGACTTTCATCTGATGATCGAAGACTTTC 1657
 1558 TTCCACATACCGAGAGAGCGCAAGGACATCACTCTCTTCAATGATCGAATTTTGA 1617
 1658 TTTACACAAAGTTGAGATGAGAGAACTTAATGTGCTCAATGACCGCAACAGCCAA 1717
 1618 TTTATGACCCGAGAGCAAAACATGAGAACATTTGGGACATCTGAGATACCGGTCAAG 1677
 1718 TTCTGTTCAATGCGCAAGGCTTGAATGCTGTCAAGAACTTAAACGAGATCTGTGAGT 1777
 1678 CCTCTTCTCATGAGCAAGCTTCAAGAGGTTGAAGAAATTAACGAGGCTGTGTGAAGCTT 1737
 1778 GCGGCAAGAACCAAGTTGCGTGAAGTTGGCTAAC---TGTAGTTGTAGTGTGATA 1834
 1738 TTGCTAATGCGCTAAGCTTGAAGAGAGCTGTGAACCTTGTGTGTGCTCCGGTATCA 1797
 1835 GCGCAAGAAATCTAAGATTTGAGAGAGAGCTGAAATGAGAAATGTTTGAAGCTGA 1894
 1798 ATGTCAACAGATTCAGAGAGAGAGAGATCGCGGATATGAGAAATGATGAATCA 1857
 1895 TCGAAGATTAAGATTTGAAGCGGCAATTCAGATGATATCATCTAATTAAGAACAAATCC 1954
 1858 TCAAGACCAACAACTTTGTTGGGCGAGTTCCGCTGATCTTGGCCAGCAACAGAGGCC 1917
 1955 GAAATGTTGAACCTTACAGATATTTGCGACAGAAAGTGTCTTGTATACAGCTGTGAT 2014
 1918 GTAAAGGAGAGCTTAATGCTAATGCTGATGCTGATACCAATGATGCTTTGTAAGCGG 1977
 2015 TGTATGAAGCTTGTGATGACATTTGTGAGGCAATGATTCGCGTTTGCACATTTG 2074
 1978 TCTATGAAGCTTGTGATGACATTTGTGAGGCAATGATTCGCGTTTGCACATTTG 2037
 2075 CAACCTGTAAGCTGTGAGACAGCGGATTAATTTGTCAATGAGGAAATCTGTTTGA 2134
 2038 GAGCGCTCAATGAGAGTCACTGATGATTAAGAGATGAGGAGTGTGCGGCTTCAATTTG 2097
 2135 ATCTTACATGATGATCAAGCTGCTGATATCTGTGATTTCTTTGAAAGTGTAGA 2194

Db 2098 ACCGTCACACCCGAAACAGCGCTGTAATCTGATGGCCGACTCTTCTGACCGGTGCAAGC 2157
QY 2195 AAGATCCATCTCACTGGATGAAGATCTCCCAAGAGGCTTGAACGAATAGAGGAAAGT 2254
Db 2158 AAGACCCAGATCACTGGGTGAAATATCTGAGACAGGGCTGACAGGCATATACGAAGT 2217
QY 2255 ATACATGGAAGATTATCTCGAGAGACTATTGACCTTGAACAGAGTGTATGATTTCTGGA 2314
Db 2218 ACACATGGAAGATTATCTCAAGAGAGTTGATGACACTGGCCGGGGTCTACGGTTCTGGA 2277
QY 2315 AGCATGTTTCAACCTTGAAGCCGTGAGAGTCGTCTTACCTTGAATGTTTATGCTC 2374
Db 2278 AGTACGTGCAAGCTTCAGAGGCGGAGACGAGCGCTACCTTGAGATGTTCTACATAC 2337
QY 2375 TTAAGTACCTGAAGCTGCTGATCATCTTCATTGGCAGAGATGA 2421
Db 2338 TGAAGTCCGCGAGCTGCGCAAGACCGTGCCTTGCAATTGACCA 2384

RESULT 4
CD726832/c 1015 bp mRNA linear EST 26-JUN-2003
LOCUS EST028 Cucurbita pepo testa subcloned cDNA Cucurbita pepo cDNA
DEFINITION clone SUST similar to Sucrose synthase, mRNA sequence.
ACCESSION CD726832
VERSION CD726832.1 GI:32277679
KEYWORDS EST.
SOURCE Cucurbita pepo
ORGANISM Cucurbita pepo
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.

REFERENCE 1 (bases 1 to 1015)
Bezold, T.N., Mathews, D., Loy, J.B. and Minocha, S.C.
TITLE Molecular analysis of the hull-less seed trait in pumpkin:
AUTHORS Expression profiles of the hull-less related genes during development
JOURNAL Unpublished (2003)
COMMENT Contact: Subhash Minocha, Dennis Mathews, Brent Loy, Todd Bezold
Dr. Minocha
University of New Hampshire
Rudman Hall, Durham, NH 03824, USA
Tel: 603 862 3840
Fax: 603 862 3784
Email: sminocha@cisunix.unh.edu

DEGENERATE primers and Tag were used to amplify cDNA for TOPO TA
(Invitrogen, Carlsbad, CA) cloning. Sequencing was performed three
times using the Dyanamic ET Terminator Sequencing kit (Amersham
pharmacia Biotech Inc, Piscataway, NJ). Sequences were identified
by NCBI BLAST(X).

FEATURES
source
1..1015
Location/Qualifiers
/organism="Cucurbita pepo"
/mol_type="mRNA"
/db_xref="taxon:3663"
/clone="SUST"
/dev_stage="20 days post-anthesis"
/clone_lib="Cucurbita pepo testa subcloned cDNA"
/note="Organ: Testa; Total RNA was isolated from 20 day
post-anthesis testa tissue and used in a subtraction
hybridization procedure as according to the Clontech
PCR-Select cDNA Subtraction kit. (PT1117-1) (Clontech, Palo
Alto, CA)."

ORIGIN

Query Match 26.5%; Score 696.6; DB 14; Length 1015;
Best Local Similarity 80.8%; Pred.No.1.3e-120;
Matches 823; Conservative 1; Mismatches 191; Indels 3; Gaps 1;
QY 563 AGGCAAGACATGATGTTGAATGACAGAAATTCAGAACTTGATCTCTTCAACATGTTT 622
Db 1015 AGGCAAGACATGATGTTGAATGATAGAAATTCGAACTTGAATCCTTCAACATGTC 956

QY 623 TGAGAAAGCAGAGAGATATCTTGGTACCTCACTCTCTGAGACACATATGTCGAATTGC 682
Db 955 TGAGAAAGCAGAGAGATATCTTGAATCACTTGACCTTGCAACCTCACTCGAGTTTG 896
QY 683 AAGACCGGTTCCAGGAATCGGTTTGGAAAGAGGTTGGGGTGACACCGCAGAAAGCGTGC 742
Db 895 CGCTGAAGTTCCAGGAATGAGCTTCCAGAGAGAGGTGGGAGACACCGCCAGAGCTGAC 836
QY 743 TCGAATGATCCAACTCTCTTGGATCTTCTGAGGCACTGATCTTGGACCTTGGAGA 802
Db 835 TTGAATGATCCAACTCTTGGATCTCTCGAGGCCCCCTGATCTTGGACCTCGAGA 776
QY 803 AGTCTTTGGAGAAATCCCATGCTGTTCAATGTTTGATCTCACTCCCAAGGATCT 862
Db 775 AGTCTTTGGAGAAATCCCATGCTTCTTAAAGTTGATCTTCTGCGCCACGTTACT 716
QY 863 TCGCTCAAGCAATGTTTGGGGTATCCCGACACCGGAGGAGGTTTGAATCTTGG 922
Db 715 TTGCAAGAGATTAAGCTCTTGGGTTACCCCGACACCGGAGGAGGTTTGAATCTTGG 656
QY 923 ATCAAGTCCGAGCTTTGGAGAAATGATGCTCTCCGTATTAAGCAACAGAACTCAACA 982
Db 655 ACCAAGTTCCGCTTTGGAGATGAATAAGCTTCAACGTATTAAGCAACAGAACTTGGACA 596
QY 983 TCACCCCTCGAAATCTCTATTTACTAGTCTTCTCTGAGTCTGTGGAAACAATGCG 1042
Db 595 TAACTCTCTGATTTCTCAATTAACACAGACTCTCTCCAGAGAGTGAAGAACGACTGCA 536
QY 1043 GTCAAGCACTTGAAAGATATACGGAACAGAGCACTCGATATCTTCAAGTACCTTCA 1102
Db 535 CACAAAGTCTCGAAGAAAGTATACGGAACAGATTAACGCCACATCTTCAAGTACCTTCA 476
QY 1103 GAACAGAAAGGAAATGTTTGAAGATGATCTCAAGATTTGAAAGAGTCTGGCCATCT 1162
Db 475 GAAATGAGAAAGGATGCTCGAAGAGTGAATCTCAAGATTCG---AAGTCTGGCCTTACC 419
QY 1163 TGGAAACTTACACAGAGATGTTGCTCATGAATCTCCAAAGAGTGAACGACGCGAC 1222
Db 418 TAGAGACTTACACCGAGATGTTTACACAGAACTCTGCAAGAAATGTCAGGAAAGCCCG 359
QY 1223 ATCTGATCATCGAAGAAACACAGCGAGCAATATGTCGCTCTCTGTCGACATTAAT 1282
Db 358 ATCTGATCATCGAAGAAACACAGCGAGCAATATGTCGCTCTCTGTCGACATTAAC 299
QY 1283 TAGGTGTACACAGTGAACCATTCGCCCATGCTTTGGAGAGACAAATATTCAGATTGAG 1342
Db 298 TAGAGATTACACAGTGAATGTTGCTCAAGCTTGGAGAAACCAATATACCTGATTGAG 239
QY 1343 ATATCTATTGGAAGAAAGCTTGAAGACAAATTCATTTCTCTTGGCAATTTACAGTGATC 1402
Db 238 ATATCTATTGGAAGAAATTTGAGAGATTAAGTATCACTTCTCATGTCMAATTCACAGTGATC 179
QY 1403 TTTTGGCAATGAACCATATACAGATTTTCATCATCAACAGATTTTCCAGAAATTTGAGAGAA 1462
Db 178 TTATTTGCATTAACCATTAACGATTTTCATCATCAAGATTTTCCAGAAATTTGAGAGAA 119
QY 1463 GCAAGACACTGTTGTCATATACAGAGACCACTGCTTTGACTCTTCTGCTGCTTACAC 1522
Db 118 GCAAGATATCTGTTGACAAATATGAGAGTCACTGSCCTTCACTCTCCGGGTCTTACAC 59
QY 1523 GTGTTGATCATGATGATGATGTTTGAATCCCAATTTCAATGTTTCCCTGATGC 1580
Db 58 GGGTGTCTCAAGCAATGACGCTGTTGATCCCAATTTCAATGTTTCACTTGAAC 1

RESULT 5
LOCUS BQ406651
DEFINITION BQ406651 689 bp mRNA linear EST 22-MAY-2002
ACCESSION GA_E00097B05f Gossypium arboreum 7-10 dpa fiber library Gossypium
BQ406651
VERSION BQ406651.1 GI:21094338
KEYWORDS EST.

Source	Organism	Reference Authors	Title	Journal Comment	Features Source
Gossypium arboreum	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Malvales; Malvaceae; Malvoideae; Gossypium.	Wang, R.A., Friesch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A.	An integrated analysis of the genetics, development, and evolution of the cotton fiber	Unpublished (2000)	1. .689 Location/Qualifiers
Gossypium arboreum				Clemson University Genomics Institute	/organism="Gossypium arboreum"
				100 Jordan Hall, Clemson, SC 29634, USA	/mol_type="rRNA"
				Tel: 864 656 7288	/strain="AKA"
				Fax: 864 656 4293	/cultivar="8400"
				Email: rwing@clemson.edu	/db_xref="taxon:29729"
				Total High Quality bases = 565	/clone="GA_Ed0097B05f"
				Seq primer: TAAATGACATCATATAGG	/cruise_type="Fibers isolated from bolls harvested 7-10 dpa"
				High quality sequence start: 2	/lab_host="E. coli"
				High quality sequence stop: 677.	/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
					/note="Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"
Query Match	25.8%; Score 677.8; DB 13; Length 689;				
Best Local Similarity	99.0%; Pred. No. 5,2e-117;				
Matches 682; Conservative	0; Mismatches 7; Indels 0; Gaps 0;				
1644	GATCGAAGACCTTCTTTACACCAAGTTGAGATGAGATGAGAACCTATGTGTCATAGA				
1	GATCGAAGACCTTCTTTACAGCAAGTTGAGATGAGATGAGAACCTATGTGTCATAGA				
1704	CCGCAACGAAGCAATTCGTTCACACATGCCAAGGCTGATGTCGACGAACCTTAACCGG				
61	CCGCAACGAAGCAATTCGTTCACACATGCCAAGGCTGATGTCGACGAACCTTAACCGG				
1764	ACTGTCGAGTGTGTCGCGCAAGAACCCAAAGTTGCGTGAATGCTACTCTGATGTTGT				
121	ACTGTCGAGTGTGTCGCGCAAGAACCCAAAGTTGCGTGAATGCTACTCTGATGTTGT				
1824	AGGTGTGATAGCGCAAGAAAGAAATCTTAAGATTTGGAAGAGAAAGCTGAATGAAGAAAT				
181	AGGTGTGATAGCGCAAGAAAGAAATCTTAAGATTTGGAAGAGAAAGCTGAATGAAGAAAT				
1884	GTTTGAGCTGATCCGACAGTACCAATTTGAACGCGCCCAATTCAGATGGAATATCTCAAT				
241	GTTTGAGCTGATCCGACAGTACCAATTTGAACGCGCCCAATTCAGATGGAATATCTCAAT				
1944	GAACAGAAATCCGAATGTTGGAACCTTAACCGATACATTTGCGACAGAAAGGTGCTTTGT				
301	GAACAGAAATCCGAATGTTGGAACCTTAACCGATACATTTGCGACAGAAAGGTGCTTTGT				
2004	ACAGCTTGATTTGATGAAGCTTTGGAATGACAGTTGTGAGAGCAATGACTTGGGTTTT				
361	ACAGCTTGATTTGATGAAGCTTTGGAATGACAGTTGTGAGAGCAATGACTTGGGTTTT				
2064	GCAACATTTGCAACCTGTAAACGCTGACACAGCGCAAGATTATGTCATGAGGAATCTGG				
421	GCAACATTTGCAACCTGTAAACGCTGACACAGCTGAGATTATGTCATGAGGAATCTGG				

QY	2124	TTTCAACATTGATCCTTACCATGGTATGATCAAGCTGCGACATATCTGTCATTTCTTTGA	2183
Db	481	TTTCAACATTGATCCTTACCATGGTATGATCAAGCTGCGACATATCTGTCATTTCTTTGA	540
QY	2184	AAAGTGTAGAAAGATCCATCTCACTGGGATTAAGATCTCCCAAGAGGCTTGAACGAAT	2243
Db	541	AAAGTGTAGAAAGATCCATCTCACTGGGATTAAGATCTCCCAAGAGGCTTGAACGAAT	600
QY	2244	AGAGGAGAGTATTCATGGAAGATTTACTGGAGAGACTATTGACCTTGACAGAGTGT	2303
Db	601	AGAGGAGAGTATTCATGGAAGATTTACTGGAGAGACTATTGACCTTGACAGAGTGT	660
QY	2304	TGGAATTCGGAAGCAGTTCCTCAACCTTG	2332
Db	661	TGGAATTCGGAAGCAGTTCCTCAACCTTG	689
RESULT 6			
LOCUS	BG439960		
DEFINITION	GA_Ea0005G07f Gossypium arboreum 7-10 dpa fiber library Gossypium		
ACCESSION	BG439960		
VERSION	BG439960.1		
KEYWORDS	EST.		
SOURCE	Gossypium arboreum		
ORGANISM	Gossypium arboreum		
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.		
AUTHORS	1 (bases 1 to 690)		
TITLE	Wing,R.A., Friesch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.		
JOURNAL	An integrated analysis of the genetics, development, and evolution of the cotton fiber		
COMMENT	Unpublished (2000)		
	Contact: Wing RA		
	Clemson University Genomics Institute		
	Clemson University		
	100 Jordan Hall, Clemson, SC 29634, USA		
	Tel: 864 656 7288		
	Fax: 864 656 4293		
	Email: twing@clemson.edu		
	Seq primer: TAAATGCACTCACTATAGCG		
	High quality sequence stop: 684.		
FEATURES			
source	1..690		
	/organism="Gossypium arboreum"		
	/mol_type="mRNA"		
	/strain="AKA"		
	/cultivar="8400"		
	/db_xref="taxon:29729"		
	/clone="GA_Ea0005G07f"		
	/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"		
	/lab_host="E. coli"		
	/clone_lib="Gossypium arboreum 7-10 dpa fiber library"		
	/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"		
ORIGIN			
Query Match	25.8%; Score 677.2; DB 12; Length 690;		
Best Local Similarity	98.8%; Pred.No. 6,7e-117;		
Matches 682; Conservative 0; Mismatches 8; Indels 0; Gaps 0;			
QY	1326	AAATATCCGATTCAGATATCTTATGGAAGAGCTTGAAGACAAATACATTTCTCTTG	1385
Db	1	AAATATTCGATTCAGATATCTTATGGAAGAGCTTGAAGACAAATACATTTCTCTTG	60
QY	1386	CCAATTACAGCTGATCTTTTGGCAATGAACCATACAGATTTTCATCATCAACGATACCTT	1445
Db	61	CCAATTACAGCTGATCTTTTGGCAATGAACCATACAGATTTTCATCATCAACGATACCTT	120
QY	1446	CCAGAAATTCGAGAAACGACGACTGTTGGTCAATACGAGAGCCACACTGCTTTCAAC	1505

Db 121 CCAGAAATTCAGAGAACAGAGACACTGTGGTCAATACAGATATCCACACTGCTTTAC 180
QY 1506 TCTTCCTGCTCTACACCGTGTGTATCATGTATCGATGTTGTTGATCCCAATTCACAT 1565
Db 181 TCTTCCTGCTCTACACCGTGTGTATCATGTATCGATGTTGTTGATCCCAATTCACAT 240
QY 1566 TGTTCCTGCTGCTGTATGATGAGATATCTCCCTTACACCGAAGAAAGCGAGTT 1625
Db 241 TGTTCCTGCTGCTGTATGATGAGATATCTCCCTTACACCGAAGAAAGCGAGTT 300
QY 1626 GAAGCATTTCCATCCTGATGATGAGATGAGATCTTCTTACACCAAGTTGATGAGAAACA 1685
Db 301 GAAGCATTTCCATCCTGATGATGAGATGAGATCTTCTTACACCAAGTTGATGAGAAACA 360
QY 1686 CTTATGTGTGCTCAATGACCGCAACCAATCTGTGTCAATGCTCAAGGCTTGATCG 1745
Db 361 CTTATGTGTGCTCAATGACCGCAACCAATCTGTGTCAATGCTCAAGGCTTGATCG 420
QY 1746 TGTCAAGAACTTAAACCGGACTGCTGTGAGTGTGCGGCAAGAACCCAAAGTTGCGTGA 1805
Db 421 TGTCAAGAACTTAAACCGGACTGCTGTGAGTGTGCGGCAAGAACCCAAAGTTGCGTGA 480
QY 1806 GCGTAACTCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1865
Db 481 GCGTAACTCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 1866 GCGTAACTCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1925
Db 541 GCGTAACTCTGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 1926 ATGATATCATCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1985
Db 601 ATGATATCATCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 1986 CACGAAAGTGTGCTTGTACAGCTTGCAAT 2015
Db 661 CACGAAAGTGTGCTTGTACAGCTTGCAAT 690

RESULT 7
BQ405173 683 bp mRNA linear EST 22-MAY-2002
LOCUS BQ405173 Ed0079E08f Gossypium arboreum 7-10 dpa fiber library Gossypium
DEFINITION GA_BQ405173 Ed0079E08f, mRNA sequence.
ACCESSION BQ405173
VERSION BQ405173.1 GI:21092860
KEYWORDS EST.
SOURCE Gossypium arboreum
ORGANISM Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 683)
Wing,R.A., Friesch,D., Yu,Y., Main,D., Rambo,T., Simmons,J.,
Henry,D., Wood,T.C., Leslie,A., and Wilkins,T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
JOURNAL COMMENT
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total High Quality bases = 573
Seq primer: TAATACGACTCACTATAGG
High quality sequence start: 5
High quality sequence stop: 669.
Location/Qualifiers
1..683
/organism="Gossypium arboreum"

FEATURES
Source 1..683
/organism="Gossypium arboreum"

/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_BQ405173"
/issue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/clone_id="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

Query Match 25.7%; Score 674; DB 13; Length 683;
Best Local Similarity 99.1%; Pred. No. 2.7e-116;
Matches 677; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 180 GAAGTGTGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 239
Db 1 GAAGTGTGCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY 240 GCGTCAATGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 299
Db 61 GCGTCAATGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 300 GAATGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 359
Db 121 GAATGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 360 AGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 419
Db 181 AGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 420 CAATCATCATTCCTCCCGCCCACTCTTCAAAATTCATGATGATGATGATGATGATGATGAT 479
Db 241 CAATCATCATTCCTCCCGCCCACTCTTCAAAATTCATGATGATGATGATGATGATGATGAT 300
QY 480 AAATGTGACCTTGTGCGCAAAATGTTTCATGATGATGATGATGATGATGATGATGATGAT 539
Db 301 AAATGTGACCTTGTGCGCAAAATGTTTCATGATGATGATGATGATGATGATGATGATGAT 360
QY 540 ATTCTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 599
Db 361 ATTCTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 600 CTTGAATGCTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
Db 421 CTTGAATGCTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 660 TGAACACCATGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719
Db 481 TGAACACCATGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 720 GGGGACACCGCAAGACCGGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
Db 541 GGGGACACCGCAAGACCGGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 780 AACTGATCTTGTGACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839
Db 601 AACTGATCTTGTGACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 840 GATTCTCACTCCCAAGGATGAT 862
Db 661 GATTCTCACTCCCAAGGATGAT 883

RESULT 8
CD486446 746 bp mRNA linear EST 01-JUN-2003
LOCUS CD486446
DEFINITION CRH2.2F08 Cotton Root and Hypocotyl lambda ZIPLOX library (CRH)
Gossypium hirsutum cDNA clone CRH2.2F08 5' similar to sucrose
synthase, mRNA sequence.
ACCESSION CD486446
VERSION CD486446.1 GI:31407411

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

EST.
Gossypium hirsutum (upland cotton)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 746)
Dowd, C., Wilson, I., and McFadden, H.
Different Gene Expression Responses in Cotton Root and Hypocotyl
tissues during infection with Fusarium Wilt Disease
Unpublished (2003)
Contact: Caltriona Dowd, Helen McFadden
Commonwealth Scientific and Industrial Research Organisation
Division of Plant Industries, Cnr Clunies Ross Street & Barry Drive,
Black Mountain, Canberra, ACT, 2601, Australia
Tel: 61 2 6246 4914, 6246 5377
Fax: 61 2 6246 5000
Email: Caltriona.Dowd@csiro.au, Helen.McFadden@csiro.au
Vector clipped sequences Bases 1-17 (GTGACCCACGCTCCG): Sali
adaptec
Seq primer: M13 reverse primer
High quality sequence stop: 746.
Location/Qualifiers

FEATURES
SOURCE

1..746
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/cultivar="DeltaEMERALD"
/db_xref="taxon:3635"
/clone="CRH2.2F08"
/tissue_type="Root and hypocotyl tissues"
/dev_stage="5 day old seedlings"
/lab_host="Y1090(ZL)"
/clone_lib="Cotton Root and Hypocotyl Lambda ZIPLOX
Library (CRH)"
/note="Vector: Lambda ZIPLOX; Site 1: Sali; Site 2: NotI;
mRNA was prepared from root and hypocotyl tissues of the
cotton cultivar DeltaEMERALD. cDNA was synthesised from a
NotI-oligo primer/adaptor using the manufacturer's
protocols (Life Technologies) and then ligated to a Sali
adaptor to facilitate directional cloning. The cDNA was
cloned into the Sali and NotI sites of the Lambda ZIPLOX
phage vector (Life Technologies). Constructed by Caltriona
Dowd and Helen McFadden."

ORIGIN

Query Match 25.3%; Score 664.8; DB 14; Length 746;
Best Local Similarity 97.2%; Pred. No. 1.4e-114;
Matches 697; Conservative 0; Mismatches 18; Indels 2; Gaps 2;

1553 CCAAAATTCACATTTTCCCTGCTGCTGATATGAGATATTAATCTTCCCTTACACCGAAG 1612
32 CCCAAATTAACATTTTCCCTGCTGCTGATATGAGATATTAATCTTCCCTTACACCGAAG 91

1613 AGAAGCGGAGGTGAGATGATTCCTCCGAGATGAGAGACCTTCTTACACCAAGTTG 1672
92 AGAAGCGGAGGTGAGATGATTCCTCCGAGATGAGAGACCTTCTTACACCAAGTTG 151

1673 AGAATGAGAACACTTATGTGTCTCAATGACCGCAACAGCAATTCGTTCACATGTC 1732
152 AGAATGAGAACACTTATGTGTCTCAATGACCGCAACAGCAATTCGTTCACATGTC 211

1733 CAAGCTTGATCGTTCAGAACTTAACCGGACTGTCGAGTGTGCGGCAAGAACCA 1792
212 CAAGCTTGATCGTTCAGAACTTAACCGGACTGTCGAGTGTGCGGCAAGAACCA 271

1793 AGTTCGCGAGTTGGCTACCTCGTAGTTGAGTGTGAGTATGGGGAAGAAATCTAAG 1852
272 AGTTCGCGAGTTGGCTACCTCGTAGTTGAGTGTGAGTATGGGGAAGAAATCTAAG 331

1853 ATTGGAAGAGAGGCTGAATGAGAAATGTTTGGCTGATCGACAAGTCAACTTGA 1912
332 ATTGGAAGAGAGGCTGAATGAGAAATGTTTGGCTGATCGACAAGTCAACTTGA 391

1913 ACGCCAAATTCAGATGATATCATCTCAATGAGACAGATCCGAATGTTGAACCTTACC 1972
392 ACGCCAAATTCAGATGATATCATCTCAATGAGACAGATCCGAATGTTGAACCTTACC 451

1973 GATACATTTTGGAGACAGAAAGAGTCCCTTTGTACAGCTGATTTGATGAGCCCTTTGAT 2032
452 GATACATTTTGGAGACAGAAAGAGTCCCTTTGTACAGCTGATTTGATGAGCCCTTTGAT 511

2033 TGACAGTTTGTGAGAGCAATGATCTTGGGTTTGGCCAAATTCGCAACTGTATACCGTGAC 2092
512 TGACAGTTTGTGAGAGCAATGATCTTGGGTTTGGCCAAATTCGCAACTGTATACCGTGAC 571

2093 CAGCCGAGATTAATGTCATCGAGAAATCTGTTTCAACATTTGATCTTACCATGATGC 2152
572 CAGCCGAGATTAATGTCATCGAGAAATCTGTTTCAACATTTGATCTTACCATGATGC 631

2153 AAGCTGCTGACATCTGTCGATTTCTTTGAAAAGTAAAGAAAGATTCATCTGACGCG 2212
632 AAGCTGCTGACATCTGTCGATTTCTTTGAAAAGTAAAGAAAGATTCATCTGACGCG 690

2213 ATTAAGATCTCCCAAGAGGCTTGAAGCAATGAGAGAGATATACATGAAAGTTT 2269
691 ATTAAGATCTCCCAAGAGGCTTGAAGCAATGAGAGAGATATACATGAAAGTTT 746

RESULT 9
BQ406766 682 bp mRNA linear EST 22-MAY-2002
LOCUS
DEFINITION
GA_Ed0098E07f Gossypium arboreum 7-10 dpa fiber library Gossypium
BQ406766
ACCESSION
BQ406766.1 GI:21094453
VERSION
BQ406766.1 GI:21094453
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 682)
Wing, R.A., Friesch, D., Yu, Y., Main, D., Rambo, T., Simmons, J.,
Henry, D., Wood, T.C., Leslie, A., and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rtwing@clemson.edu
Local High Quality bases = 571
Seq primer: TAATACGACTCACTATAGG
High quality sequence stop: 676.
Location/Qualifiers

FEATURES
SOURCE

1..682
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ed0098E07f"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 24.7%; Score 649.6; DB 13; Length 682;
Best Local Similarity 97.7%; Pred. No. 1e-111;
Matches 668; Conservative 0; Mismatches 13; Indels 3; Gaps 1;

QY 804 GTTCTTGGGAGATCCCATGCTTCAATGTTGATTTCTACTCCCAACGATCTT 863
DB 1 GTTCTTGGGAGATCCCATGCTTCAATGTTGATTTCTACTCCCAACGATCTT 60
QY 864 CGCTCAACAAATGTTTGGGGTATCCGACACCGGTGGGAGGTTGTTTCACTTTGA 923
DB 61 CGCTCAACAAATGTTTGGGGTATCCGACACCGGTGGGAGGTTGTTTCACTTTGA 120
QY 924 TCAAGTCCGAGCTTGGAGAAATGAGATGCTCTCCGTAATAAGCAACAGACTCAACAT 983
DB 121 TCAAGTCCGAGCTTGGAGAAATGAGATGCTCTCCGTAATAAGCAACAGACTCAACAT 180
QY 984 CACCCCTCGAATCCTCATTTATTACTAGACTTTCTTCTGATGCTGCGAACAACATGCGG 1043
DB 181 CACCCCTCGAATCCTCATTTATTACTAGACTTTCTTCTGATGCTGCGAACAACATGCGG 240
QY 1044 TCAACGACTTGAAGAAATGATCCGAAACAGAGACTCGAATATCTTCCAGTACCTTTAG 1103
DB 241 TCAACGACTTGAAGAAATGATCCGAAACAGAGACTCGAATATCTTCCAGTACCTTTAG 300
QY 1104 AACAGAAAGGGAATGTTCCGAAATGATCTCAAGATTGGAATAAGCTGGCCATACTT 1163
DB 301 AACAGAAAGGGAATGTTCCGAAATGATCTCAAGATTGGAATAAGCTGGCCATACTT 357
QY 1164 GGAACCTTACACAGAGATGTTGCTCATGAAATCTCCAAAGATTGCAACGCGACCA 1223
DB 358 GGAACCTTACACAGAGATGTTGCTCATGAAATCTCCAAAGATTGCAACGCGACCA 417
QY 1224 TCTGATCATCGGAACACAGAGGCAATTCCTCCCTCTTCTGCTGCAACATAAT 1283
DB 418 TCTGATCATCGGAACACAGAGGCAATTCCTCCCTCTTCTGCTGCAACATAAT 477
QY 1284 AGGTGTCAACAGTGCACCATGCCCCATGCTTTGGAGAAAGCAAAATATCCAGATTGCA 1343
DB 478 AGGTGTCAACAGTGCACCATGCCCCATGCTTTGGAGAAAGCAAAATATCCAGATTGCA 537
QY 1344 TATCTATTGGAGAAAGCTTTGAAGCAAAATACCATTTCTCTTGGCAATTTACAGCTGATCT 1403
DB 538 TATCTATTGGAGAAAGCTTTGAAGCAAAATACCATTTCTCTTGGCAATTTACAGCTGATCT 597
QY 1404 TTTTGGCAATGAAACCATAGATTTGATCATGACGATCTTTCCGGAATTTGACGAGAG 1463
DB 598 TTTTGGCAATGAAACCATAGATTTGATCATGACGATCTTTCCGGAATTTGACGAGAG 657
QY 1464 CAAGGACACTGTTGGTCAATACGA 1487
DB 658 CAAGGACACTGTTGGTCAATACGA 681

RESULT 10
CF514627 951 bp mRNA linear EST 09-SEP-2003
LOCUS Cabud0005.1IF_H03 Vitis vinifera cv. cabernet sauvignon (Clone 8)
DEFINITION Bud - CABUD Vitis vinifera cDNA clone Cabud0005.1IF_H03 5', mRNA
Sequence.
ACCESSION CF514627
VERSION CF514627.1 GI:34546395
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
REFERENCE 1 (bases 1 to 951)
AUTHORS Goes da Silva, F., Iandolo, A., Lim, H., Baek, J., Jones, K. and
Cook, D.
TITLE Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon',
berries at various developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology

One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6861
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Seq primer: ACGGTACCGCATATGACC.
Location/Qualifiers
1. 951
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon (Clone 8)"
/db_xref="taxon:29760"
/clone="Cabud0005.1IF_H03"
/sex="hermaphrodite"
/dev_stage="Pre-bloom (10-11 days before bloom)"
/lab_host="DHSalpa"
/clone_1id="Vitis vinifera cv. cabernet sauvignon (Clone 8) Bud - CABUD"
/notes="Organ: Bud; Vector: pDNR; Site 1: Sfil; Site 2: Sfil; CABUD is a cDNA library of Vitis vinifera cv. 'Cabernet Sauvignon' Clone 8 dissected buds. Samples were collected May 13, 2002 from pre-bloom plants (10-11 days before bloom), pre-veraison. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
5'-AAGCAGGTATCAACGACGAGTGGCCATTAAGCCGGG-3' and
5'-ATCTAGAGCGCGAGCGCGCATG-dT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN
Query Match 24.1%; Score 633; DB 14; Length 951;
Best Local Similarity 79.7%; Pred. No. 1,2e-108;
Matches 760; Conservative 0; Mismatches 191; Indels 3; Gaps 1;
QY 719 GGGGTGACACCGGAGAACGGCTGCTCGAGATGATCCAACTCCTTTGGATCTTTGAG 778
DB 1 GGGGTGATGACTGCTGACGCTGTCTGAGATGATTCATCTTTGGACCTTTGAG 60
QY 779 CAATGATGCTTGGACCTTGGAGAAAGTTCTTGGGAGAAATCCCATGGTGTCAATGTTG 838
DB 61 CTCCTGACCCCTGACACTTGAAGCAATTTCTTGGAGAAATCCCATGGTGTCAATGTTG 120
QY 839 TGATTTCACTCCCGACGATCTTGCCTCAAGCAATGTTTGGGTATCCGACACG 898
DB 121 TTATTTCTCCCCCATGCTACTTGTCTGAGCAATGTTCTGGGCTACCTGACACTG 180
QY 899 GTGGCAGGTTGTTAATCTTGAATCAAGTCGAGCTTTGGAGAAAGATGCTCTCC 958
DB 181 GTGGCAGGTTGTTAATCTTGAATCAAGTTCTGTCATGAGACATGAGATGCTCTCC 240
QY 959 GTATTAAGCAACAGACTCAATCAACCCCTCGAATCTCATATTACTAGACTTCTTC 1018
DB 241 GTATCAAGCAACAGACTTGAATCACTCCCAAGATCATATTGTAAGTGAATCTTCC 300
QY 1019 CTGATGCTGTGGAACAACATGCGGTCAACGACTTGAAGAAAGTATACGAAACAGACT 1078
DB 301 CTGATGCGGTAGGAGACACTTGAACACGAGCTATCGAGAAAGTTATGGAACAGACTT 360
QY 1079 CGGATATTCTTCCGATACCTTCCGAAACAGAAAGGAATTTGTCGAAATGATTCGA 1138
DB 361 CAATCATCTTTCGAGTTCTTGAACATGAGAAAGGAATTTGTCACAAATGAGATCTCA 420
QY 1139 GATTGAAAAAGTTGGCACTTATGGAACCTTACACAGAGATGTTGCTCATGAAATCT 1198
DB 421 GATTG---AAGTGTGCCATCTTGAAGAACTTACATGAGAGATGTTGCAAAAGACTTG 477
QY 1199 CCAAGAGTTGCAACGAGCAGCAGATGATGATCGAAMCAGCAGCAGCAGCAATATCG 1258
DB 478 CTACAGAGCTACAGACCAAGCAGATTTTATCATTTGCAATTAACAGATGAGAAACATTTG 537
QY 1259 TCGCTCTTGTGCTGCAATTAATTAAGTGTCAACAGTGCACCATGCGCCATGCTTTGG 1318

Db	538	TTGGCTCTTTCCTGGCTCATTAAGCTTAGGGGTTACAGAGTGCACCACTAACTATGCTCCCTGG	597
QY	1319	AGAAAGCAAAATATCCAGATTCAAGTATCTATTGGAGAAAGCTTGAAGACAAATACCTT	1378
Db	598	AGAAACCAATATATCCGAATCATGACACTTATTTGGAGAAATCTAGAGACAAATGACACT	657
QY	1379	TCTCTTGGCAATTTTACAGCTGATCTTTTGGCAATGAAACCATACAGATTTCAATCACCA	1438
Db	658	TCTCATGCGAATTTACAGCTGATCTTATTTGGCAATGAAACCATACGATTTTATCATCACCA	717
QY	1439	GTACTTTCAGAGAAATTGCAGAGAACAGACACTGTTGGTCAATACAGAGACCAACTG	1498
Db	718	GTACCTTCCAAAGAGATTCTGGAAGCAAGACACTGTTGGCAATACAGAGATCAATCTG	777
QY	1499	CTTTCACCTCTTCCTGGTCTTACCGTGTGTGAACATGATGATGTGTTGATCCCAAT	1558
Db	778	GATTCACCATCTCCGGGCTGTACAGAGATGATCCATGATGATGATGTTTGAACCCCAAT	837
QY	1559	TCAACATTGTTTCCCTGGTGTCTGATATGAGAGATTAATTTCCCTTACACGAGAGAGAC	1618
Db	838	TCAACATTGTTTACACGAGAGACAGATTAAGACATCTACTTCTCTACACTGAGAGAGAT	897
QY	1619	GGAGGTGAACATTTCCATCCTGAGATCGAAGACCTCTTACACCAAGTTG	1672
Db	898	TGAGGCTGAAAGCCCTTCATCCAGAAATTTGAAGACTCTTCTTACGCCCGGTG	951

LOCUS	DEFINITION	656 bp	mRNA	linear	EST 22-MAY-2002
BQ409552	GA_E80021E04r	Gossypium	arbitrarium	7-10 dpa	fiber library
BQ409553/c	arbitrarium	cdna	clone	GA_E80021E04r	mRNA sequence.

ACCESSION	BQ409552
VERSION	BQ409552.1
KEYWORDS	EST.
SOURCE	<i>Gossypium arboreum</i>
ORGANISM	<i>Gossypium arboreum</i>

REFERENCE
AUTHORS

1 (bases 1 to 656)
Wang, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J.,
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosoid; euroside II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 656)

TITLE	JOURNAL	COMMENT
Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A. An integrated analysis of the genetics, development, and evolution of the cotton fiber Unpublished (2000)		Contact: Wing RA

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel.: 864 656 7288
Fax: 864 656 4293
Email: rwins@clemson.edu
Total High Quality bases = 551
Seq primer: TAATCAGACTCATTAGGG
High quality sequence start: 3.
High quality sequence stop: 615.

FEATURES	Location/Qualifiers
source	1. .656
	/external: "Coccyzus leucorhous"

```

/organism="Gossypium hirsutum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ed0021E04r"
/tissue_type="Fibers isolated from bolls harvested 7-10

```

```

dpa"
/lab host="E. coli"
/clone_idb="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"

```

Query Match	24.0%;	Score 630.2;	DB 13;	Length 656;
Best Local Similarity	98.6%;	Pred. No. 4.4e-108;		
Matches 646;	Conservative 0;	Mismatches 8;	Indels 1;	Gaps 1;

QY	1972	CGATACATTTTGCACACGAAAGGTG - CTTTGTGACAGCTTGACATTTGATGAGCCCTTTGG	2030
Db	656	CGATACATTTTGCACACGAAAGGTGCCCTTTGTGACAGCCGTGACATTTGATGAGCCCTTTGG	597
QY	2031	ATTGACAGTTGTGGAGGCATGACTTGCGGTTTGGCCAACTTGGCAACTGTGACACTGTGACGGTGG	2090
Db	596	ATTGACAGTTGTGGAGGCATGACTTTGCGGTTTGGCCAACTTGGCAACTGTGACACTGTGACGGTGG	537
QY	2091	ACCAGCCGAGATTAATTTGTCATGGAATCTGGTTTCAACATTTGATCCTTAACATGGTGA	2150
Db	536	ACCAGCTGAGTTATTTGTCATGGAATCTGGTTTCAACATTTGATCCTTAACATGGTGA	477
QY	2151	TCAAGCTGCTGACATCTACTGTCGATTTTCTTTGAAAAGTGTGAAGAAATCATCTGACTG	2210
Db	476	TCAAGCTGCTGACATCTACTGTCGATTTTCTTTGAAAAGTGTGAAGAAATCATCTGACTG	417
QY	2211	GGATTAAGATCTCCCAAGAGGCTTGAACGAAATAGAGAGAGATACATGAAAGATTTA	2270
Db	416	GGATTAAGATCTCCCAAGAGGCTTGAACGAAATAGAGAGAGATACATGAAAGATTTA	357
QY	2271	CTCGAGAGACTATATTGACCCTTGACAGAGTGTATGATTTCTGGAAGCATGTTTCCAACCT	2330
Db	356	CTCGAGAGACTATATTGACCCTTGACAGAGTGTATGATTTCTGGAAGCATGTTTCCAACCT	297
QY	2331	TGAACGCCGAGAGAGTGTGCTGTTACCTTTGAGATGTTTTATGCTCTTAAAGTACCGTAAGCT	2390
Db	296	TGAACGCCGAGAGAGTGTGCTGTTACCTTTGAGATGTTTTATGCTCTTAAAGTACCGTAAGCT	237
QY	2391	GGCTGAATCAAGTTCATTTGACAGAGAGTAATTTGAACCTGTGTAATPAACATTGGGCCGG	2450
Db	236	GGCTGAATCAAGTTCATTTGACAGAGAGTAATTTGAAGCTGTGTAATPAACATTGGGCCGG	177
QY	2451	TTTTCTTGGAGAAATAATATCTGTTTTGTGTAATTTCAATGGAGAACTCCTTTGTATTT	2510
Db	176	TTTTCTTGGAGAAATAATATCTGTTTTGTGTAATTTCAATGGAGAACTCCTTTGTATTT	117
QY	2511	CATCTGTCTTTTCTTTTCTTTTTCGCGGACATTTGTGAACATGGGGTTTGCGCC	2570
Db	116	CATCTGTCTTTTCTTTTCTTTTTCGCGGACATTTGTGAACATGGGGTTTGCGCC	57
QY	2571	CGTCAATTCAGTTAAATATGTGTACTTTGTTTTCAAAAAAAAAAAAAAAAAAAAAA 2625	
Db	56	CGTCAATTCAGTTAAATATGTGTATTTGTTTTCAAAAAAAAAAAAAAAAAAAAAA 2	

RESULT	12
BQ415686/c	
LOCUS	BQ415686
DEFINITION	637 bp mRNA linear EST 22-MAY-2002 GA_EB0102A07r Gossypium arboreum 7-10 dpa fiber library Gossypium arboreum cDNA clone GA_EB0102A07r, mRNA sequence.

ACCESSION	BQ415686
VERSION	BQ415686.1
KEYWORDS	GI:21103373
SOURCE	EST.
	<i>Gossypium arboreum</i>

ORGANISM
Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Pharmetophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosales; Malvaceae; Malvaceae

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 637)
Wang, R.A., Fritsch, D., Yu, Y., Main, D., Rambo, T., Simmons, J.,
Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber

JOURNAL COMMENT
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288

Fax: 864 656 4293
Email: rtwing@clmson.edu
Total High Quality bases = 539
Seq primer: TAATGACTCATTATAGCG
High quality sequence start: 3
High quality sequence stop: 632.
Location/Qualifiers

FEATURES

source

1. 637
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ed0102A07r"
/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"
/lab_host="E. coli"
/clone_id="Gossypium arboreum 7-10 dpa fiber library"
/note="Vector: pBR-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 23.7%; Score 623.2; DB 13; Length 637;

Best Local Similarity 98.7%; Pred. No. 9.2e-107;
Matches 628; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1989 GAAAGTGCCTTTGACAGCTGATGTGATGAGCCTTTGGATTGACAGTTGTGAGGC 2048
636 GAAAGTGCCTTTGACAGCTGATGTGATGAGCCTTTGGATTGACAGTTGTGAGGC 577
2049 AATGCTGCGGTTGCCAACAATTGGCAACGCTGTAACGCTGACACGCGAGATTATGT 2108
576 AATGCTGCGGTTGCCAACAATTGGCAACGCTGTAACGCTGACAGATTATGT 517
2109 CCATGGAAATCTGCTTCAACATTCCTTACCATGGTATCAAGCTGCTGACATCT 2168
516 CCATGGAAATCTGCTTCAACATTCCTTACCATGGTATCAAGCTGCTGACATCT 457
2169 CGTCGATTTCTTTGAAAAGTGAAGAAAGATCCATCTCACTGGGATTAAGATCTCCAAAG 2228
456 CGTCGATTTCTTTGAAAAGTGAAGAAAGATCCATCTCACTGGGATTAAGATCTCCAAAG 397
2229 AGGCTTGAAGCAAGAGAGAGAAATATATCATGGAAGATTAATCTGGAGAGATTAAGAC 2288
396 AGGCTTGAAGCAAGAGAGAGAAATATATCATGGAAGATTAATCTGGAGAGATTAAGAC 337
2289 CCTGACAGAGATGATGATCTGAGAACATGTTTCCAACCTTGAAACCCCTGAGAGTCG 2348
336 CCTGACAGAGATGATGATCTGAGAACATGTTTCCAACCTTGAAACCCCTGAGAGTCG 277
2349 TCGTTACCTTGAGATGTTTATGCTCTTAAGTACCGTAAGCTGCTGAATCAATTCATT 2408
276 TCGTTACCTTGAGATGTTTATGCTCTTAAGTACCGTAAGCTGCTGAATCAATTCATT 217
2409 GGCAGAGAGATTAATGAAACCTGTAATTAACATGAGGCGCGTTTCTTGGAGAAATAT 2468
216 GGCAGAGAGATTAATGAAACCTGTAATTAACATGAGGCGCGTTTCTTGGAGAAATAT 157
2469 ATTCTGTTTGAATTTCAATTTGAGAAAGCTCTTGTATTTCACTGTCTTTTCTTT 2528
156 ATTCTGTTTGAATTTCAATTTGAGAAAGCTCTTGTATTTCACTGTCTTTTCTTT 97
2529 TCGTTTTCGCGGCAATTTGTTGAACATGAGGCTTGCAGCCCTGCAATTCAGTTAAAT 2588
96 TCGTTTTCGCGGCAATTTGTTGAACATGAGGCTTGCAGCCCTGCAATTCAGTTAAAT 37
2589 ATGTGACTTTTGTCTTTTCAAAAAAATTTTTTTTTT 2624
36 ATGTGACTTTTGTCTTTTCAAAAAAATTTTTTTTTT 1

RESULT 13
AI731292 AT1731292 664 bp mRNA linear EST 11-JUN-1999
LOCUS

DEFINITION BNHG19102 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar
to (U73588) sucrose synthase (Gossypium hirsutum), mRNA sequence.
ACCESSION AI731292
VERSION AI731292.1 GI:5050144
KEYWORDS EST.
SOURCE Gossypium hirsutum (upland cotton)
ORGANISM Gossypium hirsutum

REFERENCE
AUTHORS Blewett M., Marx E.C., Davy D.F. and Burr B.
TITLE 1 (bases 1 to 664)
JOURNAL Ests from developing cotton fiber
COMMENT Unpublished (1999)
Contact: Ben Burr
Biology Department
Brookhaven National Laboratory
Upton, NY 11973, USA
Tel: 516-344-3396
Fax: 516-344-3407
Email: burr@bnl.bnl.gov

FEATURES

source
Location/Qualifiers

1. 664
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/cultivar="Acala Maxxa"
/db_xref="taxon:3635"
/tissue_type="Immature fiber"
/dev_stage="Six days post anthesis"
/lab_host="XLI-Blue"
/clone_id="Six-day Cotton fiber"
/note="Vector: pBluescript II KS+"

ORIGIN

Query Match 23.7%; Score 623.2; DB 9; Length 664;

Best Local Similarity 97.7%; Pred. No. 9.1e-107;
Matches 642; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

1713 GCCAATCTCTGTTCAATATGCCAAGCTTGATCTGTCAAGAACTTAACCGACTCGTGA 1772
1 GCCAATCTCTGTTCAATATGCCAAGCTTGATCTGTCAAGAACTTAACCGACTCGTGA 60
1773 GTGTCGGCGCAAGAACCCAAAGTTCGTGAGTTGGCTAACCTCGTATGTTAGTGTGA 1832
61 GTGTCGGCGCAAGAACCCAAAGTTCGTGAGTTGGCTAACCTCGTATGTTAGTGTGA 120
1833 TAGCGGAAGGAATCTTAAGATTTTGGAGAGAGAGGCTGAATGAAGAAATGTTAGCT 1892
121 TAGCGGAAGGAATCTTAAGATTTTGGAGAGAGAGGCTGAATGAAGAAATGTTAGCT 180
1893 GATGACAAGTACAATCTTGAAACGCGCAATTCAGATGATATCATCTCAAAATGAACAGAT 1952
181 GATGACAAGTACAATCTTGAAACGCGCAATTCAGATGATATCATCTCAAAATGAACAGAT 240
1953 CCGAAATGTTGAACCTTACCGATATCTTTGGAGACAGAAAGTGCTTTGTACAGCTGC 2012
241 CCGAAATGTTGAACCTTACCGATATCTTTGGAGACAGAAAGTGCTTTGTACAGCTGC 300
2013 ATTGTATGAAGCCTTGTGATTTGACAGTTGTGGAGGCAATGACTGGGTTGCCAATTT 2072
301 ATTGTATGAAGCCTTGTGATTTGACAGTTGTGGAGGCAATGACTGGGTTGCCAATTT 360
2073 CGCAACTGTATACGCTGAGACCGAGCAATTAATGTCATGAGGAATCTGTTTCAACAT 2132
361 CGCAACTGTATACGCTGAGACCGAGCAATTAATGTCATGAGGAATCTGTTTCAACAT 420
2133 TGATCTTACCATGATGATCAAGCTGTGACATATCTGTCGATTTCTTTGAAAAGTATA 2192
421 TGATCTTACCATGATGATCAAGCTGTGACATATCTGTCGATTTCTTTGAAAAGTATA 480
2193 GAAAGATCCATCTCACTGGGATTAAGATCTCCAAAGAGGCTTGAAGAAATAGAGGAGA 2252

Db 481 GAAAGATCCATTCACCTGGGATTAAGATCTCCCAAGAGGCTTGAACGATTAAGAGAGA 540

Qy 2253 GTATACATGAGAGATTACTCGGAGAGACTATTGACCCCTGACAGAGGTGTATGATTCG 2312

Db 541 ATATACATGAGAGATTACTCGGAGAGACTATTGACCCCTGACAGAGGTGTATGATTCG 600

Qy 2313 GAGCATGTTT-CCAACTTGAAGCCCGTGAAGTGTGTGTTACCTTGAGATGTTT 2368

Db 601 GAAACATGTTTCCCACTTGAAACCCGTAAGAGTGTGTGTTACCTTGAAATGTTT 657

RESULT 14

LOCUS BG442235 642 bp mRNA linear EST 15-MAR-2001

DEFINITION GA_Ea001E08f Gossypium arboreum 7-10 dpa fiber library Gossypium

ACCESSION BG442235

VERSION BG442235.1 GI:13351887

KEYWORDS EST.

SOURCE Gossypium arboreum

ORGANISM Gossypium arboreum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium. 1 (bases 1 to 642)

AUTHORS Wing,R.A., Friesch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.

TITLE An integrated analysis of the genetics, development, and evolution of the cotton fiber

JOURNAL Unpublished (2000)

COMMENT Contact: Wing RA

Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7286

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: TAATACGACTCCTATGAGG

High quality sequence stop: 640.

FEATURES

Source

1. 642

/organism="Gossypium arboreum"

/mol_type="mRNA"

/strain="AKA"

/cultivar="8400"

/db_xref="taxon:29729"

/clone="GA_Ea001E08f"

/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"

/lab_host="E. coli"

/clone_idb="Gossypium arboreum 7-10 dpa fiber library"

/note="Vector: PBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 23.7%; Score 623; DB 12; Length 642;

Best Local Similarity 99.1%; Pred. No. 1e-106;

Matches 637; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1916 GCCAATTCAGATGATTCATCTCAATGAAGACAGATCCGAATGTGTAACCTTACCGAT 1975

Db 1 GCCAATTCAGATGATTCATCTCAATGAAGACAGATCCG-AATGTGTAACCTTACCGAT 59

Qy 1976 ACATTGGGACAGAAAGAGTCCCTTGTACAGCCCTGCACTTGTATGAAGCCTTTGATTTGA 2035

Db 60 ACATTGGGACAGAAAGAGTCCCTTGTACAGCCCTGCACTTGTATGAAGCCTTTGATTTGA 119

Qy 2036 CAGTTTGAGAGCAATGACTTGGGTTGCCAACAATTCGCAACCTGTAAACGTTGACAG 2095

Db 120 CAGTTTGAGAGCAATGACTTGGGTTGCCAACAATTCGCAACCTGTAAACGTTGACAG 179

Qy 2096 CCGAGATTATTTGCTCCAGGAAATCTGGTTTCAACATTGATCTTACCAGGTATCAAG 2155

Db 180 CTGAGATTATTTGCTCCAGGAAATCTGGTTTCAACATTGATCTTACCAGGTATCAAG 239

Qy 2156 CTGCTGACATCTGCTCGATTCTTTGAAAAGTGAAGAGATCATCTCACTGGATA 2215

Db 240 CTGCTGACATCTGCTCGATTCTTTGAAAAGTGAAGAGATCATCTCACTGGATA 299

Qy 2216 AGATCTCCCAAGAGGCTTGAACGATTAAGAGAGAACTATACATGAAATTACTCGG 2275

Db 300 AGATCTCCCAAGAGGCTTGAACGATTAAGAGAGAACTATACATGAAATTACTCGG 359

Qy 2276 AGAGCTATTGACCTGACAGAGTGTATGATTTCTGGAACGATTTCCAACTTGAAC 2335

Db 360 AGAGCTATTGACCTGACAGAGTGTATGATTTCTGGAACGATTTCCAACTTGAAC 419

Qy 2336 GCCGTGAGAGTGTGTGTTACCTTGAGATGTTTATGCTTTAAGACCGTAGCTGCTG 2395

Db 420 GCCGTGAGAGTGTGTGTTACCTTGAGATGTTTATGCTTTAAGACCGTAGCTGCTG 479

Qy 2396 AATCAGTTCCATTGGCAGAGAGATTAATTGAACCTGTTAAATTAACATTGGCCGGTTT 2455

Db 480 AATCAGTTCCATTGGCAGAGAGATTAATTGAACCTGTTAAATTAACATTGGCCGGTTT 539

Qy 2456 CTGGAGAAATTAATTTCTGTTTGTATTTCAATTGAGAACTCTTTGATTTATCT 2515

Db 540 CTGGAGAAATTAATTTCTGTTTGTATTTCAATTGAGAACTCTTTGATTTATCT 599

Qy 2516 TGTCTTTTCCCTTTCCCTTTTTCGCCGCGCATTTGTTGAACATG 2558

Db 600 TGTCTTTTCCCTTTCCCTTTTTCGCCGCGCATTTGTTGAACATG 642

RESULT 15

LOCUS BE054845 871 bp mRNA linear EST 07-MAR-2001

DEFINITION GA_Ea0032D04f Gossypium arboreum 7-10 dpa fiber library Gossypium

ACCESSION BE054845

VERSION BE054845.2 GI:13245895

KEYWORDS EST.

SOURCE Gossypium arboreum

ORGANISM Gossypium arboreum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium. 1 (bases 1 to 871)

AUTHORS Wing,R.A., Friesch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry,D., Wood,T.C., Leslie,A. and Wilkins,T.A.

TITLE An integrated analysis of the genetics, development, and evolution of the cotton fiber

JOURNAL Unpublished (2000)

COMMENT On Jun 8, 2000 this sequence version replaced gi:8381998.

Contact: Wing RA

Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7286

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: TAATACGACTCCTATGAGG

High quality sequence stop: 581.

FEATURES

Source

1. 871

/organism="Gossypium arboreum"

/mol_type="mRNA"

/strain="AKA"

/cultivar="8400"

/db_xref="taxon:29729"

/clone="GA_Ea0032D04f"

/tissue_type="Fibers isolated from bolls harvested 7-10 dpa"

/lab_host="E. coli"

/clone_idb="Gossypium arboreum 7-10 dpa fiber library"

/note="Vector: PBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 23.6%; Score 620.2; DB 10; Length 871;
Best Local Similarity 86.0%; Pred. No. 3e-106;
Matches 734; Conservative 0; Mismatches 113; Indels 6; Gaps 4;

```
QY 1460 GAAGGAAGACACCTGGTGGTGAATACGAGAGCCACACCTGCTTCACTCTTCCGTGCTCT 1519
DB 1 GAAGCAAGACACCTGGTGGTGAATACGAGAGCCACACCTGCTTCACTCTTCCGTGCTCT 60
QY 1520 ACCGTTGTACATGATGATTCGATGTTGATCCCAATTCAACATGTTTCCCTGGTG 1579
DB 61 ACCGTTGTACATGATGATTCGATGTTGATCCCAATTCAACATGTTTCCCTGGTG 120
QY 1580 CTGATATGAGATATATCTTCCCTTACACCGAAGAGAGCGAGGTTGAAGCATTTTCCATC 1639
DB 121 CTGATATGAGATATATCTTCCCTTACACCGAAGAGAGCGAGGTTGAAGCATTTTCCATC 180
QY 1640 CTGAGATGGAAGACCTTCTTTACACCAAGTTGAGAAATGAAGAACATTATGTGTCTCA 1699
DB 181 CTGAGATGGAAGACCTTCTTTACACCAAGTTGAGAAATGAAGAACATTATGTGTCTCA 240
QY 1700 ATGACCGCAACAAAGCCAAATTCGTCACAAATGCCAAGCTTGATCGTGTCAAGAACTTAA 1759
DB 241 ATGACCGCAACAAAGCCAAATTCGTCACAAATGCCAAGCTTGATCGTGTCAAGAACTTAA 300
QY 1760 CCGGACTCGTGGAGTGTGCGGCAAGAACCCAAAGTTGCGTGAATCTGCTAACTCTGTAG 1819
DB 301 CCGGACTCGTGGAGTGTGCGGCAAGAACCCAAAGTTGCGTGAATCTGCTAACTCTGTAG 360
QY 1820 TTGTAGTGTGTATATGCGGAAAGAAATCTAAAGATTTGGAAGAGAAAGCTGAATGAAGA 1879
DB 361 TTGTAGTGTGTATATGCGGAAAGAAATCTAAAGATTTGGAAGAGAAAGCTGAATGAAGA 420
QY 1880 AAATGTTGAGCTGATCGACAAAGTACAACTTGAACGGCCAAATTCAGATGATATCATCTC 1939
DB 421 AAATGTTGAGCTGATCGACAAAGTACAACTTGAACGGCCAAATTCAGATGATATCATCTC 480
QY 1940 AAATGAACAGAAATCCGAAATGTTGAACCTTTACCGATACATTTGCGACACGAAAGGTGCT 1999
DB 481 AAATGAACAGAAATCCGAAATGTTGAACCTTTACCGATACATTTGCGACACGAAAGGTGCT 539
QY 2000 TTGTACAGCTGATGATGATGAAGCTTTGATTTGACAGTTGTGAGGCAATGACTTGGC 2059
DB 540 TTGTACAGCTGATGATGATGAAGCTTTGATTTGACAGTTGTGAGGCAATGACTTGGC 599
QY 2060 GTTTCGCAACATTC-GCAACCTGTACGCTGAGCCAGCCGAGATTATGTCATGGGAAA 2118
DB 600 GATTGGCCACATTCAGCAACTGTGACCGAGGACCAAGCTGGAATATGACCATGGGAAA 659
QY 2119 TCTGGTTTCACATTTGATCTTTACCATGCTGATCAAGCTGTCACATCTGCTGATTTTC 2178
DB 660 TGTGGTTTAAACATGGGTTCTTTACATGGGATGAGAAAGTAAATTTCCCAAGAAATG 719
QY 2179 TTTGAAAAGTGTAAAGAAATGATCATCTCATGGAATA--AGATCTCCCAAGAGGCTTG 2235
DB 720 TTTTAAAAAGGTGTGAAAAATTCATGAGCTTGAATATATATCAACAAAGAGGCTGG 779
QY 2236 AAACGAATAGAGAGAGATATCATGAAATTTACT-CGAGAGACTATTGACCTGAC 2294
DB 780 AGCGAGTAAATATTAATAATTCATAGCAGAGGTTCTAGGAAAACACATAGAACCCGAC 839
QY 2295 AGGAGTGTATGA 2307
DB 840 AGGAGAGAGGGA 852
```

Search completed: June 5, 2004, 08:30:34
Job time : 9938 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model1

Run on: June 2, 2004, 14:46:39 ; Search time 61 Seconds
(without alignments)
3733.333 Million cell updates/sec

Title: US-10-003-405-2
4247
Sequence: 1 MAEALTRVHSIRERLEDTL.....EMFYALKRKLAESVPLAE 806

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4245	100.0	806	AAU97898	AAU97898 Cotton su
2	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
3	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
4	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
5	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
6	3643.5	85.8	805	AAU97898	AAU97898 Cotton su
7	3603.5	84.8	808	AAU97898	AAU97898 Cotton su
8	3295	77.6	816	AAU97898	AAU97898 Cotton su
9	3273	77.1	816	AAU97898	AAU97898 Cotton su
10	3273	77.1	816	AAU97898	AAU97898 Cotton su
11	3270.5	77.0	815	AAU97898	AAU97898 Cotton su
12	3268.5	77.0	815	AAU97898	AAU97898 Cotton su
13	3267	76.9	814	AAU97898	AAU97898 Cotton su
14	3259.5	76.7	808	AAU97898	AAU97898 Cotton su
15	3254.5	76.6	808	AAU97898	AAU97898 Cotton su
16	3250.5	76.5	808	AAU97898	AAU97898 Cotton su
17	3248	76.4	808	AAU97898	AAU97898 Cotton su
18	3246.5	76.4	808	AAU97898	AAU97898 Cotton su
19	3234.5	76.2	808	AAU97898	AAU97898 Cotton su
20	3170.5	74.7	773	AAU97898	AAU97898 Cotton su
21	3092.5	72.8	773	AAU97898	AAU97898 Cotton su
22	3069.5	72.3	777	AAU97898	AAU97898 Cotton su
23	3069.5	72.3	777	AAU97898	AAU97898 Cotton su
24	2972	70.0	809	AAU97898	AAU97898 Cotton su
25	2962	69.7	809	AAU97898	AAU97898 Cotton su

26	2961	69.7	802	AAE28499	AAE28499 Corn suc
27	2875	67.7	797	ABR39586	ABR39586 A. thalia
28	2866.5	67.5	805	ABR39582	ABR39582 A. thalia
29	2866.5	67.5	805	ABR39585	ABR39585 A. thalia
30	2841	66.9	766	AAE66222	AAE66222 Sucrose-s
31	2312.5	54.5	51.4	ABR91573	ABR91573 Herbicide
32	2184.5	51.4	786	ADC07862	ADC07862 Rice prot
33	2147.5	50.6	798	ADC07860	ADC07860 Rice prot
34	2133	50.2	843	ABR93633	ABR93633 Herbicide
35	1747	41.1	806	AAW53103	AAW53103 Anabaena
36	1440	33.9	514	ADC07854	ADC07854 Rice prot
37	1383	32.6	395	ABG69054	ABG69054 Amino aci
38	1174.5	27.7	348	AAU16313	AAU16313 Pinus rad
39	879	20.7	225	ABG69052	ABG69052 Amino aci
40	707	16.6	198	AAU95667	AAU95667 Plant col
41	698	16.4	204	ABW73686	ABW73686 DNA clone
42	643.5	15.2	149	AAU16309	AAU16309 Eucalyptu
43	621	14.6	217	AAU16312	AAU16312 Pinus rad
44	603	14.2	242	ABG69053	ABG69053 Amino aci
45	574.5	13.5	158	AAU16284	AAU16284 Pinus rad

ALIGNMENTS

RESULT 1	AAU97898	AAU97898 standard; protein, 806 AA.
AC	AAU97898;	
DT	13-AUG-2002 (first entry)	
DE	Cotton sucrose synthase Susy protein.	
KW	Cotton; Sucrose synthase; Susy; fibre; seed; transgenic; plant; enzyme.	
OS	Gossypium hirsutum.	
FH	Key	Location/Qualifiers
FT	Misc-difference 414	/label= Unknown
FT		/note= "Encoded by NAC"
XX	WO200245485-A1.	
PD	13-JUN-2002.	
PF	07-DEC-2001; 2001WO-AU001580.	
PR	08-DEC-2000; 2000AU-00001975.	
PR	08-DEC-2000; 2000US-0251852P.	
PA	(CSIR) COMMONWEALTH SCI & IND RES ORG.	
XX	Llewellyn D, Furbank R, Ruan Y;	
DR	WPI; 2002-463779/49.	
XX	N-PSDB; ABX52880.	
PT	Altering fibre development or properties of a fibre producing plant by	
PT	modulating sucrose synthase activity and/or expression in such plants,	
PT	useful for enhancing fibre yield and quality and for increasing seed	
XX	size.	
PS	Claim 5; Page 57-60; 62pp; English.	
CC	This invention relates to a novel method for altering fibre development	
CC	or properties of a fibre producing plant by modulating sucrose synthase	
CC	(Susy) activity and/or expression in such plants. The invention also	
CC	comprises a fibre producing plant comprising a chimaeric gene in its	
CC	genome, the seeds of the plant and fibre isolated from the plant. The	
CC	method is useful for altering fibre development or properties of a fibre	

CC producing plant like cotton plant. Therefore, the method is useful for
 CC enhancing fibre yield, enhancing fibre quality and for increasing seed
 CC size in a fibre producing plant. The present sequence represents the
 CC cotton sucrose synthase (Susy) protein used to create the transgenic
 CC plant of the invention

XX Sequence 806 AA;

Query Match 100.0%; Score 4245; DB 5; Length 806;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAERALTRVSHSLRERLDETLAHRNEIALLSRLEGKGILOHQHIIIEFAIPEENRK 60
 Db 1 MAERALTRVSHSLRERLDETLAHRNEIALLSRLEGKGILOHQHIIIEFAIPEENRK 60
 Qy 61 KLANGAFPEVLKASOEAVLPWPVALAVPRPGWEYIRVVNVALVVEELVAAYLHFKE 120
 Db 61 KLANGAFPEVLKASOEAVLPWPVALAVPRPGWEYIRVVNVALVVEELVAAYLHFKE 120
 Qy 121 ELVVGSSNGNFVLELDEPFNSPPRPYLSKISGVGFLLNRHSAKLFDHKSMPHLE 180
 Db 121 ELVVGSSNGNFVLELDEPFNSPPRPYLSKISGVGFLLNRHSAKLFDHKSMPHLE 180
 Qy 181 FLRVHCHGKMMNDRIQNLALQHVLRKAEEYLGTLPPETPCAEFEHRFOEIGLERGW 240
 Db 181 FLRVHCHGKMMNDRIQNLALQHVLRKAEEYLGTLPPETPCAEFEHRFOEIGLERGW 240
 Qy 241 GDTAERVLEMIQLLLDLLEATDPTLEKFLGRIMVFNVVVLTTHGYAQNVLGYPTG 300
 Db 241 GDTAERVLEMIQLLLDLLEATDPTLEKFLGRIMVFNVVVLTTHGYAQNVLGYPTG 300
 Qy 301 GQVYVYLDQVRLAENEMLRIRKQOGINTPRILITRLLPDVAGTTCGRLKXYGTEHS 360
 Db 301 GQVYVYLDQVRLAENEMLRIRKQOGINTPRILITRLLPDVAGTTCGRLKXYGTEHS 360
 Qy 361 DILRVPTREKQIVRKWISREKWPYLETYTEDVAHESIKELHGTDDLIGNKSQGNIV 420
 Db 361 DILRVPTREKQIVRKWISREKWPYLETYTEDVAHESIKELHGTDDLIGNKSQGNIV 420
 Qy 421 ASLAAHKLGVTOCTIAHLEKTKYPPDSDIYWKKEDEKHFSCOFADLFAMNHDTFIITS 480
 Db 421 ASLAAHKLGVTOCTIAHLEKTKYPPDSDIYWKKEDEKHFSCOFADLFAMNHDTFIITS 480
 Qy 481 TFOEIASGKDTVGOYESHFAFLPGLVYVVGIDVDFDKFNIVSPGADMEIYFYTEBK 540
 Db 481 TFOEIASGKDTVGOYESHFAFLPGLVYVVGIDVDFDKFNIVSPGADMEIYFYTEBK 540
 Qy 541 RLKGFHPEIEDLLYTKVNEEHLCVLDRNRPILFTMPRLDRVXNLTGLVEMCGNPKLR 600
 Db 541 RLKGFHPEIEDLLYTKVNEEHLCVLDRNRPILFTMPRLDRVXNLTGLVEMCGNPKLR 600
 Qy 601 ELANLVVVGGRRRKESKLEKAEKMKFELIDKYNLNGORWISSQNRIRANVELYXYI 660
 Db 601 ELANLVVVGGRRRKESKLEKAEKMKFELIDKYNLNGORWISSQNRIRANVELYXYI 660
 Qy 661 CDTGAFVOPALYEAFFGLTVVEAMTCGLPTPATCGGPAEITIVGKSGFNIDPHYGDOA 720
 Db 661 CDTGAFVOPALYEAFFGLTVVEAMTCGLPTPATCGGPAEITIVGKSGFNIDPHYGDOA 720
 Qy 721 DILVDFEKKCKDDPSHMDKISQGLKRIEKKYTWKISERLLTLTGYYGFWKHVSNNLERR 780
 Db 721 DILVDFEKKCKDDPSHMDKISQGLKRIEKKYTWKISERLLTLTGYYGFWKHVSNNLERR 780
 Qy 781 ESRRYLEMFYALKTRKLAESVPLAE 806
 Db 781 ESRRYLEMFYALKTRKLAESVPLAE 806

RESULT 2
 AAB16282
 ID AAB16282 standard; protein; 805 AA.
 XX

AC AAB16282;
 XX 31-OCT-2000 (first entry)
 XX
 DE Eucalyptus grandis sucrose synthase protein sequence SEQ ID NO:44.
 XX
 XX Eucalyptus grandis; pinus radiata; Monterey pine; modification;
 KM plant cell wall; polysaccharide biosynthetic pathway;
 XX transgenic plant.
 XX
 OS Eucalyptus grandis.
 PN MO20002092-A2.
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99NO-NZ000169.
 XX
 PR 13-OCT-1998; 98US-00170862.
 PR 11-AUG-1999; 99US-0148426P.
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 P1 Bloksberg LN;
 XX
 DR MPI; 2000-339328/29.
 DR N-PSDB; AAA67087.
 XX
 FT New genes encoding proteins involved in a plant polysaccharide
 PT biosynthetic pathway, useful for modulating or altering the
 PT polysaccharide content, composition or structure of the plant.
 PS Claim 17; Page 58-60; 301pp; English.
 XX
 CC The present invention describes isolated polynucleotides (PN) comprising
 CC a sequence selected from one of 835 nucleotide sequences given in
 CC AAA67073 to AAA67907, their (reverse) complements, sequences producing an
 CC expectation (E) value of 0.01 or less compared to the 835 sequences,
 CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
 CC 835 sequences or sequences that are degenerately equivalent or allelic to
 CC the 835 sequences. The polynucleotides are used to modify the activity of
 CC a polypeptide involved in a polysaccharide biosynthetic pathway in the
 CC plant. They are especially used to modulate or alter the polysaccharide
 CC content, composition or structure of the plant. AAB16268 to AAB16340 are
 CC proteins encoded by some of the polynucleotide sequence given in the
 CC present invention
 XX
 SQ Sequence 805 AA;
 Query Match 85.8%; Score 3643.5; DB 3; Length 805;
 Best Local Similarity 85.0%; Pred. No. 7.7e-316;
 Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;

Qy 1 MAERALTRVSHSLRERLDETLAHRNEIALLSRLEGKGILOHQHIIIEFAIPEENRK 60
 Db 1 MADRMLTRSHSLRERLDETLAHRNDIVAFSLRVEAKGKGILOHQHIIIEFAIPEENRK 60
 Qy 61 KLANGAFPEVLKASOEAVLPWPVALAVPRPGWEYIRVVNVALVVEELVAAYLHFKE 120
 Db 61 KLANGAFPEVLKASOEAVLPWPVALAVPRPGWEYIRVVNVALVVEELVAAYLHFKE 120
 Qy 121 ELVVGSSNGNFVLELDEPFNSPPRPYLSKISGVGFLLNRHSAKLFDHKSMPHLE 180
 Db 121 ELVVGSSNGNFVLELDEPFNSPPRPYLSKISGVGFLLNRHSAKLFDHKSMPHLE 180
 Qy 181 FLRVHCHGKMMNDRIQNLALQHVLRKAEEYLGTLPPETPCAEFEHRFOEIGLERGW 240
 Db 181 FLRVHCHGKMMNDRIQNLALQHVLRKAEEYLGTLPPETPCAEFEHRFOEIGLERGW 240
 Qy 241 GDTAERVLEMIQLLLDLLEATDPTLEKFLGRIMVFNVVVLTTHGYAQNVLGYPTG 300
 Db 241 GDTAERVLEMIQLLLDLLEATDPTLEKFLGRIMVFNVVVLTTHGYAQNVLGYPTG 300

```

Qy 301 GGVVYILDQVRLNEMMLRIKQOGLNITPRILITRLLPDAVGTTCGGRLEKVGTEHS 360
Db 301 GGVVYILDQVRLNEMMLRIKQOGLNITPRILITRLLPDAVGTTCGGRLEKVGTEHS 360
Qy 361 DILRVFPRTKGIYKWKISREPEKWPYLETYTEVDAHEISKEHGTPLDIIIGNSDGNIV 420
Db 361 DILRVFPRTKGIYKWKISREPEKWPYLETYTEVDAHEISKEHGTPLDIIIGNSDGNIV 420
Qy 421 ASLNAHKLGVTOCTIAHALEKTKYPSDSIYWKKEDEKTHFSQCFADLFAANNHTDPIITS 480
Db 421 ASLNAHKLGVTOCTIAHALEKTKYPSDSIYWKKEDEKTHFSQCFADLFAANNHTDPIITS 480
Qy 481 TFOEIASGKDTVGVYESHMTAFTLPGLYRVVHGIDVDFPKFNIVSGADMEIYFPYTEKR 540
Db 481 TFOEIASGKDTVGVYESHMTAFTLPGLYRVVHGIDVDFPKFNIVSGADMEIYFPYTEKR 540
Qy 541 RLKGFHPRIEDLLYTKVNEEHLCVLNDNRNKPILFTMRPLDRVKNLTGLVWCGNPKLR 600
Db 541 RLKGFHPRIEDLLYTKVNEEHLCVLNDNRNKPILFTMRPLDRVKNLTGLVWCGNPKLR 600
Qy 601 ELANLVVVGDRRKESKDLSEKAEMKMFELIDKYNLNGQFRWISSQNNRIRNVELYRI 660
Db 601 ELANLVVVGDRRKESKDLSEKAEMKMFELIDKYNLNGQFRWISSQNNRIRNVELYRI 660
Qy 661 CDTGAFVQPALYEAFTGLTVEAMTCGLPTFATCNGGPAEIIYHGKSGFNIDPHYGDAA 720
Db 661 CDTGAFVQPALYEAFTGLTVEAMTCGLPTFATCNGGPAEIIYHGKSGFNIDPHYGDAA 720
Qy 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKISERLLTGLVYGFWMKVSNLERR 780
Db 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKISERLLTGLVYGFWMKVSNLERR 780
Qy 781 ESRRLYEMFYALKTRKLAESVPLAE 806
Db 781 ESRRLYEMFYALKTRKLAESVPLAE 806
Qy 780 ESRRLYEMFYALKTRKLAESVPLAE 805
Db 780 ESRRLYEMFYALKTRKLAESVPLAE 805

```

RESULT 3

AAB16336 standard; protein; 805 AA.

AAB16336;

31-OCT-2000 (first entry)

Eucalyptus grandis sucrose synthase protein sequence SEQ ID NO:144.

Eucalyptus grandis; pinus radiata; Monterey pine; modification;

plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;

transgenic plant.

Eucalyptus grandis.

MO200022092-A2.

20-APR-2000.

08-OCT-1999; 99WO-NZ000169.

13-OCT-1998; 98US-00170862.

11-AUG-1999; 99US-0148426P.

(GENE-) GENESIS RES & DEV CORP LTD.

(FLET-) FLETCHER CHALLENGE FORESTS LTD.

Blokberg LN;

WPI; 2000-339328/29.

N-PSDB; AAB67143.

New genes encoding proteins involved in a plant polysaccharide

biosynthetic pathway, useful for modulating or altering the

```

Pr polysaccharide content, composition or structure of the plant.
XX Claim 17; Page 108-110; 301pp; English.
PS The present invention describes isolated polynucleotides (PN) comprising
XX a sequence selected from one of 835 nucleotide sequences given in
CC AAB67073 to AAB67907, their (reverse) complements, sequences producing an
CC expectation (E) value of 0.01 or less compared to the 835 sequences.
CC sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the
CC 835 sequences or sequences that are degenerately equivalent or allelic to
CC the 835 sequences. The polynucleotides are used to modify the activity of
CC a polypeptide involved in a polysaccharide biosynthetic pathway in the
CC plant. They are especially used to modulate or alter the polysaccharide
CC content, composition or structure of the plant. AAB16268 to AAB16340 are
CC proteins encoded by some of the polynucleotide sequence given in the
CC present invention.
XX
SQ Sequence 805 AA;
Query Match 85.8%; Score 3643.5; DB 3; Length 805;
Beet Local Similarity 85.0%; Pred. No. 7.7e-316;
Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;
Qy 1 MAERALTREHSLRERLDELTLAHRNEIATLSRIEKGKGILOHQIILEFEALPEENRK 60
Db 1 MADRLTSHSLRERLDELTLAHRNDIYAFISRVAKKGILQHQIFAPEALISESRA 60
Qy 61 KLANGAFPEVILKASOEALVLPWVALAVRPPGWERYIRVNVHALVEELVAYELHFE 120
Db 61 KLIDGAFGEVILKSTGEALVSPWVALAVRPPGWERYIRVNVHALVLEQLVAYELHFE 120
Qy 121 ELVGSNGNVLLEDEFPNSSFPRPLSSIGNGVEFLNRHLSAKLPHDSEHPLLE 180
Db 121 ELADSLNGNVLLEDEFPNSSFPRPLSSIGNGVEFLNRHLSAKLPHDSEHPLLE 180
Qy 181 FLRVHCHGKMMMLNDRIQNLNALOHVLRKAEVYGTLPETPCADEFERFOEIGLERGM 240
Db 181 FLVHCHGKMMMLNDRIQNVFSLQHVLRKAEVYLTSLKETPYQSOFHKGQELGERGM 240
Qy 241 GDTAERVLEMIQLDLLEADPCTLEKFLRIPVFNVLTPHGYPAQDNVGLYPTG 300
Db 241 GDTAERVLEMIQLDLLEADPCTLEKFLRIPVFNVLTPHGYPAQDNVGLYPTG 300
Qy 301 GGVVYILDQVRLNEMMLRIKQOGLNITPRILITRLLPDAVGTTCGGRLEKVGTEHS 360
Db 301 GGVVYILDQVRLNEMMLRIKQOGLNITPRILITRLLPDAVGTTCGGRLEKVGTEHS 360
Qy 361 DILRVFPRTKGIYKWKISREPEKWPYLETYTEVDAHEISKEHGTPLDIIIGNSDGNIV 420
Db 361 DILRVFPRTKGIYKWKISREPEKWPYLETYTEVDAHEISKEHGTPLDIIIGNSDGNIV 420
Qy 421 ASLNAHKLGVTOCTIAHALEKTKYPSDSIYWKKEDEKTHFSQCFADLFAANNHTDPIITS 480
Db 421 ASLNAHKLGVTOCTIAHALEKTKYPSDSIYWKKEDEKTHFSQCFADLFAANNHTDPIITS 480
Qy 481 TFOEIASGKDTVGVYESHMTAFTLPGLYRVVHGIDVDFPKFNIVSGADMEIYFPYTEKR 540
Db 481 TFOEIASGKDTVGVYESHMTAFTLPGLYRVVHGIDVDFPKFNIVSGADMEIYFPYTEKR 540
Qy 541 RLKGFHPRIEDLLYTKVNEEHLCVLNDNRNKPILFTMRPLDRVKNLTGLVWCGNPKLR 600
Db 541 RLKGFHPRIEDLLYTKVNEEHLCVLNDNRNKPILFTMRPLDRVKNLTGLVWCGNPKLR 600
Qy 600 ELANLVVVGDRRKESKDLSEKAEMKMFELIDKYNLNGQFRWISSQNNRIRNVELYRI 660
Db 600 ELANLVVVGDRRKESKDLSEKAEMKMFELIDKYNLNGQFRWISSQNNRIRNVELYRI 660
Qy 661 CDTGAFVQPALYEAFTGLTVEAMTCGLPTFATCNGGPAEIIYHGKSGFNIDPHYGDAA 720
Db 661 CDTGAFVQPALYEAFTGLTVEAMTCGLPTFATCNGGPAEIIYHGKSGFNIDPHYGDAA 720
Qy 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKISERLLTGLVYGFWMKVSNLERR 780
Db 721 DILVDFEKKCKDPSHMDKISQGLKRIEKKYTWKISERLLTGLVYGFWMKVSNLERR 780

```

```
Db      720 ELVDVFNKKCIDQSHWDEISKGAMORIEEKYTKWISERLLNLTAAYGFWKHTNLD RR 779
Qy      781 ESRRYLEMFYALKYRKLAESVPLAE 806
      780 ESRRYLEMFYALKYRKPLAQSVPAVE 805

RESULT 4
AAB28141
ID      AAB28141 standard; protein; 805 AA.
AC      AAB28141;
DT      02-FEB-2001 (first entry)
DE      Sucrose synthase promoter protein #2.
KW      Promoter; eucalyptus; pine; gene transcription.
OS      Eucalyptus grandis.
EN      WO200058474-A1.
PD      05-OCT-2000.
PE      24-FEB-2000; 2000WO-NZ000018.
PR      25-MAR-1999; 99US-00276559.
PR      30-JUL-1999; 99US-0146591P.
XX      (GENE-) GENESIS RES & DEV CORP LTD.
PA      (FLET-) FLETCHER CHALLENGE FORESTS LTD.
PI      Perera R, Rice SJ, Eagleton CK;
XX      WPI; 2000-647236/62.
DR      N-PSDB; AAC62807.
XX      Novel promoter sequences useful for modulating transcription of plant DNA
PT      sequences of interest and production of polypeptides.
XX      Claim 4; Page 70-72; 93pp; English.
XX      The present invention relates to promoter sequences from eucalyptus and
CC      pine. The present protein is encoded by one such promoter. The coding
CC      sequence for the present protein is useful for modulating the
CC      transcription of DNA sequences of interest. The sequences may also be
CC      used to tag or identify an organism or its reproductive material
XX      SQ
Sequence 805 AA;

Query Match      85.8%; Score 3643.5; DB 3; Length 805;
Best Local Similarity 85.0%; Pred. No. 7,7e-316;
Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;
```

```
Db      241 GDTAERVLEMTQLLDLLEAPDCTLEKFLDRVPMFNVNIMSPHGFADDDVLGYPTDG 300
Qy      301 GQVYVILDQVRALENEMLRIRKQGINITPRILLITRLPDVAGTTGGRLEKYGTES 360
Db      301 GQVYVILDQVRALLENEMLRIRKQGLDITPRILLITRLPDVAGTTGGRLEKYGTES 360
Qy      361 DILRVPRTEKGYRKWISREKWPYLETTYEDVAHEISKELHGTDLIIGNXSDGNIY 420
Db      361 HILRVPRNEKGVARKWISRFE-VMPYLERYTEDEVASELGELOGKPDLLIIGNYSDGNIY 419
Qy      421 ASLLAHKLVTOCTIAHLEKTKYRPSDDIYWKULDEKYHSCOPTADLFANNHTDFTITS 480
Db      420 ASLLAHKLVTOCTIAHLEKTKYRPSDDIYWKULDEKYHSCOPTADLIANNHTDFTITS 479
Qy      481 TFOEIASKDTVGQYESHTAFTLPGLYRVVHGIDVDPKKNIVSPGADMEIYFYTEKR 540
Db      480 TFOEIASKDTVGQYESHNMFTLPGLYRVVHGIDVDPKKNIVSPGADMSIYFAYTEQR 539
Qy      541 RLKHFPEIDDLTYKVENEBHLCVLDRNKPILFTMPRLDRVKNLTGLYEWCGKPKLR 600
Db      540 RLKSFHPEIEELLFSDVENKEHLCVLKKDKKKPILFTMARLDRVKNLTGLYEWGKSKLR 599
Qy      601 ELANLVVGGDRRRESKDLREKAKMKPFELIDKYNLNGQFRWISQMNRIYVELRYI 660
Db      600 ELANLVVGGDRRKSQDLREKAKMKPYDLEKIKLNGQFRWISQMNRIYVELRYI 659
Qy      661 CDTKGAFVOPALYEAFGLTVVEAMTCGLPTFATCGGPAEIIYHKGSGFNIDPYGDOAA 720
Db      660 CDTKGAVVQPAIYAFGLTVVEAMTCGLPTFATCGGPAEIIYHKGSGYHNDPYHGDOAA 719
Qy      721 DIIVDFEKKCKQDSHWDKISQGLKRIIEKYTKWISERLLITGYGFKWYSNLERR 780
Db      720 ELVDVFNKKCIDQSHWDEISKGAMORIEEKYTKWISERLLNLTAAYGFWKHTNLD RR 779
Qy      781 ESRRYLEMFYALKYRKLAESVPLAE 806
Db      780 ESRRYLEMFYALKYRKPLAQSVPAVE 805

RESULT 5
AAB80759
ID      AAB80759 standard; protein; 805 AA.
XX      AAB80759;
AC      26-MAR-2002 (first entry)
DT      26-MAR-2002 (first entry)
DE      Eucalyptus grandis promoter polypeptide #10.
KW      Eucalyptus grandis promoter polypeptide #10.
XX      Promoter; pine; leaf; flower; pollen; bud; meristem; xylogenesis;
KW      temporally regulated promoter; Pinus radiata; Eucalyptus grandis.
XX      Eucalyptus grandis.
EN      WO200198485-A1.
PD      27-DEC-2001.
PE      20-JUN-2001; 2001WO-NZ000115.
PR      20-JUN-2000; 2000US-00598401.
PR      28-NOV-2000; 2000US-00724624.
XX      (GENE-) GENESIS RES & DEV CORP LTD.
PA      (FLET-) FLETCHER CHALLENGE FORESTS IND LTD.
PI      Perera R, Rice S, Eagleton C, Lasham A;
XX      WPI; 2002-114583/15.
DR      N-PSDB; ABK17072.
XX      Novel polynucleotide promoter sequences from pine and Eucalyptus useful
PT      for modifying expression of endogenous and/or heterologous
```

PT polynucleotides in transgenic plants.

XX Example 13; Page 86-88; 121pp; English.

CC The invention relates to isolated promoter sequences from *Pinus radiata*
 CC and *Eucalyptus grandis*, comprising a sequence chosen from leaf-, root-,
 CC flower-, pollen-, bud-, meristem-specific promoters or temporally
 CC regulated promoters such as xylogenesis-specific promoters. The promoter
 CC polypeptides and their related polynucleotides are useful in the
 CC production of genetic constructs, used for modifying gene expression in a
 CC target organism, in particular a plant. The method is useful for
 CC modifying expression of a polynucleotide that comprises an intron
 CC sequence, through removal of the intron sequence. The method is useful
 CC for modifying growth and development of plants, and cellular responses to
 CC external stimulus, such as environmental factors and disease pathogens.
 CC The sequences are useful in genome and physical mapping, in positional
 CC cloning of genes, in various assays to determine biological activity, to
 CC raise antibodies, to isolate corresponding interacting proteins and other
 CC compounds, and to quantitatively determine levels of interacting proteins
 CC or other compounds. Sequences AB080745-AA080764 represent *Pinus radiata*
 CC and *Eucalyptus grandis* polypeptides of the invention

XX Sequence 805 AA;

Query Match 85.8%; Score 3643.5; DB 5; Length 805;
 Best Local Similarity 85.0%; Pred. No. 7.7e-316;
 Matches 685; Conservative 61; Mismatches 59; Indels 1; Gaps 1;

QY 1 MAERALTTRVHSLRERLDETLAHRNEILALSRIEKGKGILOHQIILEFEAIPENRK 60
 DB 1 MADRMLTRSHSLRERLDETLAHRNDIVAFLSRVEAKGIIQHQIIFAEFAIESERA 60
 QY 61 KLANGAFPEVLAQSGAIVLPVVALAVPRPGWMEYIVVNVHVALVEELTYAEYLHFEK 120
 DB 61 KLIDAFGEVLKSTGEAIVSPVVALAVPRPGWMEHIVNVHVALVEELTYAEYLHFEK 120
 QY 121 ELVDSSNGNFVLELDFEFPNSFPPTLSKISGNGVEFLNHLSAKLPHDKESMHPLE 180
 DB 121 ELADSSNGNFVLELDFEFPNTSFPPTLSKISGNGVEFLNHLSAKLPHDKESMHPLE 180
 QY 181 FLRVHCHGKKNMMLNDRIQNLALQVLRKAEYLGTLPEPTPCAEFEHRFOEIGLERGW 240
 DB 181 FLQVHCYKGNMNVNARIQNVSLQHVLRKAEYLTSLKEPIPYSGFEKHPGEIGLERGW 240
 QY 241 GPTARVLEMIQLLDLLEATPCTIEKLGRIPMWFNVVILTPHYFPQDNVLYGYPDNG 300
 DB 241 GPTARVLEMIQLLDLLEATPCTIEKLGRIPMWFNVVILTPHYFPQDNVLYGYPDNG 300
 QY 301 GGVVYIILDOVRALLENMMLRIKQGGINTPRIILITRLPDVAGTTCGGRLEKVVGTSHS 360
 DB 301 GGVVYIILDOVRALLENMMLRIKQGGINTPRIILITRLPDVAGTTCGGRLEKVVGTSHS 360
 QY 361 DILRVFPRTKGIIVRKWISRFKWPYLETYEDVAHEISKEIAGTPDILIGNXSDGNIV 420
 DB 361 HILRVFPRTKGIIVRKWISRFKWPYLETYEDVAHEISKEIAGTPDILIGNXSDGNIV 420
 QY 421 ASLLAHKLGVTCTTAHALEKTKYPSDQIYMKKLEDKTHFSCQFTADLPANMHTDPIITS 480
 DB 420 ASLLAHKLGVTCTTAHALEKTKYPSDQIYMKKLEDKTHFSCQFTADLPANMHTDPIITS 480
 QY 481 TPOEAGSKDVTGQVESHHTAFTLPGLYRVVHGIDVDPFNIVSPADMEIYFPYTEERK 540
 DB 480 TPOEAGSKDVTGQVESHHTAFTLPGLYRVVHGIDVDPFNIVSPADMEIYFPYTEERK 540
 QY 541 RLKHFPRIEDLLYTKVENEHMLCVLNDNRKPIILFTMPRLDVKVNLTLGVENCGKNPKLR 600
 DB 540 RLKHFPIRIEDLLYTKVENEHMLCVLNDNRKPIILFTMPRLDVKVNLTLGVENCGKNPKLR 600
 QY 601 ELANLVVGGDRKRSKDLSEKAEKMKPELIDKYNLNGQFWRISQNNRINVELYRYT 660
 DB 600 ELANLVVGGDRKRSKDLSEKAEKMKPELIDKYNLNGQFWRISQNNRINVELYRYT 660
 QY 661 CDTKGFVQPALVEAFGLTVEAMTCGLPTPATCNGPABIIIVHGKSGNIDIPYHGDQA 720

DB 660 CDTKGFVQPALVEAFGLTVEAMTCGLPTPATCNGPABIIIVHGKSGNIDIPYHGDQA 719
 QY 721 DILVDFPEKCKDPDHPMKISQGGIKRIEKKYTWKTYSEERLLTLGVYGFWMGVNLERK 780
 DB 720 ELVDVFPFKCKIDOSHMOEISGAMQRIEKKYTWKTYSEERLLTLGVYGFWMGVNLERK 779
 QY 781 ESRRTLEMFVALKYRKLAESVPLAE 806
 DB 780 ESRRTLEMFVALKYRKLAESVPLAE 805

RESULT 6
 ABB92501
 ID ABB92501 standard; protein; 808 AA.
 AC ABB92501;
 DT 31-MAY-2002 (first entry)
 DE Herbicidally active polypeptide SEQ ID NO 1712.
 KW Herbicidal; plant; agriculture; herbicide.
 OS Arabidopsis thaliana.
 PN WO200210210-A2.
 PD 07-FEB-2002.
 PF 28-AUG-2001; 2001WO-EP009892.
 PR 28-AUG-2001; 2001WO-EP009892.
 PA (PAB) BAYER AG.
 PI Tietjen K, Weidler M;
 DR WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.

XX Claim 5; SEQ ID NO 1712; 261pp + Sequence Listing; English.

CC The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides

XX Sequence 808 AA;

Query Match 85.1%; Score 3614.5; DB 5; Length 808;
 Best Local Similarity 84.6%; Pred. No. 3e-313;
 Matches 681; Conservative 58; Mismatches 65; Indels 1; Gaps 1;

QY 2 AERALTTRVHSLRERLDETLAHRNEILALSRIEKGKGILOHQIILEFEAIPENRK 61
 DB 4 AERVITRVSQGERLDATLVAQKNEVPALLSRVEAKGIIQHQIIFAEFAIPETQK 63
 QY 62 LANGAFPEVLAQSGAIVLPVVALAVPRPGWMEYIVVNVHVALVEELTYAEYLHFEK 121
 DB 64 LKGAFFELLSAQAIVLPVVALAVPRPGWMEYIVVNVHVALVEELTYAEYLHFEK 123
 QY 122 LVDSSNGNFVLELDFEFPNSFPPTLSKISGNGVEFLNHLSAKLPHDKESMHPLEF 181

Db 124 LVDGIKGNFTLDELPEPPNNAFPPTLNTKTYIGDGEVELNRLHLSAKLFHDKESLHPLKLF 183
Qy 182 LRVHCHGKNMMLNDRIQNLALQHVLKRAEYLGTLPPETPCAFEPHRFOEIGLERGMG 241
Db 184 LRLSHGCKNTMLNRIQNLALQHVLKRAEYLMELKPELTYSPFEHKGIGLERGMG 243
Qy 242 DTAERVENMIQLLDLLEATDPTCLEKPLGRIPWPNVILTPHGYFQDNLVGYPDG 301
Db 244 DTAERVENMIQLLDLLEATDPTCLEKPLGRIPWPNVILTPHGYFQDNLVGYPDG 303
Qy 302 QVVYILDOVRALENEMLRIRKQGLNITPRILITRLPDAGTTCGRLKVGTEHSD 361
Db 304 QVVYILDOVRALETMLQRIKQGLNITPRILITRLPDAGTTCGRLKVGTEHSD 363
Qy 362 ILRVFPRTEKGIYRKWISRFKWPYLETYTEDVAHEISKELGTPDLLIGNXSDGNIVA 421
Db 364 ILRVFPRTEKGIYRKWISRFKWPYLETYTEDVAHEISKELGTPDLLIGNXSDGNIVA 422
Qy 422 SLIAHKLGVTOCTIAHALEKTYPPDSIYWKLEDPKHFSCQPTADLFAMNHTDITITST 481
Db 423 SLIAHKLGVTOCTIAHALEKTYPPDSIYWKLEDPKHFSCQPTADLFAMNHTDITITST 482
Qy 482 FOEISGSDTGVGOYESHAFPLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 541
Db 483 FOEISGSDTGVGOYESHAFPLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 542
Qy 542 LKHPHPEIEDLLYTKVENEHLCVLDNRKPIILFTMPRLDRVKNLTGLVEMCGKNPKLRE 601
Db 543 LKHPHPEIEDLLYTKVENEHLCVLDNRKPIILFTMPRLDRVKNLTGLVEMCGKNPKLRE 602
Qy 602 LANLVVGGDRRKSKDLLEKAEKMKMPELLDKYNLNGQFNMISQOMRINNVELRYIYC 661
Db 603 LANLVVGGDRRKSKDLLEKAEKMKMPELLDKYNLNGQFNMISQOMRINNVELRYIYC 662
Qy 662 DTKGAFVOPALYEAGLTVEAMTCGLPTFATCNGGPAEIIYHGSGFNIDPYHDOAAD 721
Db 663 DTKGAFVOPALYEAGLTVEAMTCGLPTFATCNGGPAEIIYHGSGFNIDPYHDOAAD 722
Qy 722 ILVDFPEKCKDPHMDKISQGLKRIEKEYTWKTYSERLLTLTGVGFWKHSVNLBRE 781
Db 723 SLADFCTCKHDPHMDKISQGLKRIEKEYTWKTYSERLLTLTGVGFWKHSVNLBRE 782
Qy 782 SRRYLEMFPYALKYRKLAESVPLAEE 806
Db 783 SRRYLEMFPYALKYRKLAESVPLAEE 807
RESULT 7
ABB93562
ID ABB93562 standard; protein; 808 AA.
AC ABB93562;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 2773.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX
PA (FARB) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX

DR WPI; 2002-269010/31.
XX Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX
PS Claim 5; SEQ ID NO 2773; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides
XX
SQ Sequence 808 AA;

Query Match 84.8%; Score 3603.5; DB 5; Length 808;
Best Local Similarity 83.9%; Pred. No. 2.9e-312;
Matches 675; Conservative 68; Mismatches 61; Indels 1; Gaps 1;

Qy 2 AERALTTRVHSLREBLDETLAHRNEILLALSRIEGKGGKGIQHQIILFEPAIPEENRK 61
Db 4 AERMTVRHSQREBLNETLVSERNEVLALLSRVAKGKGIQHQIILFEPAIPEOTRK 63
Qy 62 LANGAFPEVLKASGEALVLPWVALAVPRGQWVEYLRVYNHALVVEELTAAYLHPEE 121
Db 64 LEGGFPEQLLSTGEALVLPWVALAVPRGQWVEYLRVYNHALVVEELTAAYLHPEE 123
Qy 122 LVDSSNGNFVLALDPEFPNFPPTLSKISIGCVETFLNRLSAKLFHDKESMPLLEF 181
Db 124 LVDGKGNFTLDELPEPPNNAFPPTLNTKTYIGDGEVELNRLHLSAKLFHDKESLHPLKLF 183
Qy 182 LRVHCHGKNMMLNDRIQNLALQHVLKRAEYLGTLPPETPCAFEPHRFOEIGLERGMG 241
Db 184 LRLSHGCKNTMLNRIQNLALQHVLKRAEYLMELKPELTYSPFEHKGIGLERGMG 243
Qy 242 DTAERVENMIQLLDLLEATDPTCLEKPLGRIPWPNVILTPHGYFQDNLVGYPDG 301
Db 244 DTAERVENMIQLLDLLEATDPTCLEKPLGRIPWPNVILTPHGYFQDNLVGYPDG 303
Qy 302 QVVYILDOVRALENEMLRIRKQGLNITPRILITRLPDAGTTCGRLKVGTEHSD 361
Db 304 QVVYILDOVRALETMLQRIKQGLNITPRILITRLPDAGTTCGRLKVGTEHSD 363
Qy 362 ILRVFPRTEKGIYRKWISRFKWPYLETYTEDVAHEISKELGTPDLLIGNXSDGNIVA 421
Db 364 ILRVFPRTEKGIYRKWISRFKWPYLETYTEDVAHEISKELGTPDLLIGNXSDGNIVA 422
Qy 422 SLIAHKLGVTOCTIAHALEKTYPPDSIYWKLEDPKHFSCQPTADLFAMNHTDITITST 481
Db 423 SLIAHKLGVTOCTIAHALEKTYPPDSIYWKLEDPKHFSCQPTADLFAMNHTDITITST 482
Qy 482 FOEISGSDTGVGOYESHAFPLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 541
Db 483 FOEISGSDTGVGOYESHAFPLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEER 542
Qy 542 LKHPHPEIEDLLYTKVENEHLCVLDNRKPIILFTMPRLDRVKNLTGLVEMCGKNPKLRE 601
Db 543 LKHPHPEIEDLLYTKVENEHLCVLDNRKPIILFTMPRLDRVKNLTGLVEMCGKNPKLRE 602
Qy 602 LANLVVGGDRRKSKDLLEKAEKMKMPELLDKYNLNGQFNMISQOMRINNVELRYIYC 661
Db 603 LANLVVGGDRRKSKDLLEKAEKMKMPELLDKYNLNGQFNMISQOMRINNVELRYIYC 662
Qy 662 DTKGAFVOPALYEAGLTVEAMTCGLPTFATCNGGPAEIIYHGSGFNIDPYHDOAAD 721
Db 663 DTKGAFVOPALYEAGLTVEAMTCGLPTFATCNGGPAEIIYHGSGFNIDPYHDOAAD 722

PR 22-JUN-2001; 2001US-0300112P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 20-DEC-2001; 2001US-0342327P.
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 XX Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;
 PI Glazebrook J, Katagiri F, Kreps J, Provart N, Rieke D;
 XX
 DR MPI; 2003-229341/22.
 DR N-PSDB; ADC07855.
 XX
 PR New plant genes encoding polypeptides having an activity involved in or
 PT associated with the synthesis, metabolism or degradation of carbohydrates
 PT in the plant grain useful in generating plants having improved
 PT nutritional properties.
 XX
 PS Claim 1; SEQ ID NO 122, 130pp; English.
 XX
 CC This invention, in the area of plant biotechnology, relates to novel
 CC polynucleotides comprising a nucleotide sequence encoding a protein which
 CC is involved in or associated with the synthesis, metabolism or
 CC degradation of carbohydrates in the plant grain and the expression of
 CC which is up-regulated during grain filling. The plant is selected from
 CC corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,
 CC sugarbeet, wheat, and rice. The invention may be useful for the
 CC improvement of protein, oil, starch, fibre and moisture content of the
 CC cereal grains. In addition, carbohydrate levels may be modified to a more
 CC desirable level using the present invention. The present sequence is the
 CC amino acid sequence of a rice protein of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPD
 CC at ftp.wipd.int/pub/publihedget_sequences.
 XX
 SQ Sequence 816 AA;
 Query Match 77.1%; Score 3273; DB 7; Length 816;
 Best Local Similarity 76.1%; Pred. No. 9,9e-283;
 Matches 613; Conservative 91; Mismatches 99; Indels 2; Gaps 2;
 QY 3 ERLRVHSLRRLBETLAAHNEELALSRLEGKGLQHQTILF-EAIPENRKK 61
 DB 7 DRLSLRHSVRERIDSLSAHPNELVAVPTRLVNLGKMLQAHQIIAEVNNAISSADREK 66
 QY 62 LANGAFEEYKASOEALVLPWVALAVRPGVMEYIRVNVAVLVEELTVAEYHFKEE 121
 DB 67 LKDGAFEDVLRSAQEGIVASFWALAIRPGVMEYIRVNVAVLVEELTVAEYHFKEE 126
 QY 122 LVYDGSNGNFVLELDFEPNSSFPPTLSKSIQNGVEFLNRHLAKLPHDKESMPLLEF 181
 DB 127 LVEEGTNNNFVLELDFEPNNSFPPTLSKSIQNGVQFLNRHLSKLPDKESMPLLEF 186
 QY 182 LRVHCHGKGMNMLNRIQNLNLQVLRKAEYVLTLPETPCAFERFQSIGLERQMG 241
 DB 187 LRAHNYKGMNMLNRIQNLNLQVLRKAEYVLTLPETPCAFERFQSIGLERQMG 246
 QY 242 DTAERVLEWIOQLDLLEATDPTLEKFLGRIIPVFNVILPHGYPADNVLGPDPTGG 301
 DB 247 DCAKASQETIHLDLLEAPDSTLEKFLGRIIPVFNVILPHGYPADNVLGPDPTGG 306
 QY 302 QVVYVILDOVRALENEMLRIRKQGLNITRILIIITRLPDAVGTTCGRLERKVGTEHSD 361
 DB 307 QVVYVILDOVRAMENEMLRIRKQGLNITRILIIITRLPDAVGTTCGRLERKVGTEHSD 366
 QY 362 ILRVPRTEKGIIVRKWISFEKWPVLETYEDVAHEISKELHGPDLITGXSGNIVA 421
 DB 367 ILRVPRTEKGIIVRKWISFEKWPVLETYEDVAHEISKELHGPDLITGXSGNIVA 425
 QY 442 SLIAHLGVTGCTIIAHLAKTKYRPSDIIYKKLEDPKHSFCOFTDLPFMMNTDPIITST 481
 DB 446 CILAHMGVTHCTIIAHLAKTKYRPSDIIYKKLEDPKHSFCOFTDLPFMMNTDPIITST 485
 QY 482 FOEIASKDTVGQYESHSTAFTLPGLYRVVHGIDVEDPKENIVSPGADMEIYPPYTEKKR 541

DB 486 FOEIASKDTVGQYESHSTAFTLPGLYRVVHGIDVEDPKENIVSPGADMEIYPPYTEKKR 545
 QY 542 LKHFPRIEDLTYKVENEEHLCVLDRNRPILFTMPRLDRVKNLITGLVEMGKPKLRE 601
 DB 546 LTSIHPRIEELTYEVNNEHKPMLKDRNRPILFTMPRLDRVKNLITGLVEMGKPKLRE 605
 QY 602 LANIIVVGGRRKESKDLLEKAEKKNFELIDKYNLNGQFRWISSQNNRIRNVELYRYIC 661
 DB 606 LVNLVWVCGDHGNSKDEQAEFKMFDLEQYNLNGHIRMISAQNNRVANGELYRYIC 665
 QY 662 DTGAFVQPALYKFAFGITVWEAMTCGLPTPATCNGSBAEITVHKSQFNIDPYGDDAAD 721
 DB 666 DTGAFVQPALYKFAFGITVWEAMTCGLPTPATCNGSBAEITVHKSQFNIDPYGDDAAD 725
 QY 722 ILVDFPEKCKDPESHMDKISQGLKRIEKKYTWKIVSERLTLTGVGFWGVNLSRRE 781
 DB 726 LTVFPEKCKDPESHMDKISQGLKRIEKKYTWKIVSERLTLTGVGFWGVNLSRRE 785
 QY 782 SRRYLEMFFALKYRKLAESVPLAEE 806
 DB 786 TRRYLEMLYALKYRTMASTVPLAVE 810
 RESULT 10
 ID ADC08209 standard; protein, 816 AA.
 XX
 AC ADC08209;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Rice protein sequence Seq ID514 related to grain filling.
 XX
 KW plant biotechnology; carbohydrate synthesis; carbohydrate metabolism;
 KW carbohydrate degradation; carbohydrate; plant grain; grain filling; corn;
 KW tomato; banana; canola; cotton; peanut; sorghum; tobacco; sugarbeet;
 KW wheat; rice; protein; oil; starch; fibre; moisture content; cereal grain;
 KW gene; ds; plant.
 XX
 OS Oryza sativa.
 XX
 FN WO200300905-A2.
 XX
 PD 03-JAN-2003.
 XX
 PD 21-JUN-2002; 2002WO-1B002450.
 XX
 PR 22-JUN-2001; 2001US-0300112P.
 PR 26-SEP-2001; 2001US-0325277P.
 PR 20-DEC-2001; 2001US-0342327P.
 XX
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Zhu T, Cheng W, Briggs S, Cooper B, Goff SA, Moughamer T;
 PI Glazebrook J, Katagiri F, Kreps J, Provart N, Rieke D;
 XX
 DR MPI; 2003-229341/22.
 DR N-PSDB; ADC08208.
 XX
 PT New plant genes encoding polypeptides having an activity involved in or
 PT associated with the synthesis, metabolism or degradation of carbohydrates
 PT in the plant grain useful in generating plants having improved
 PT nutritional properties.
 XX
 PS Claim 34; SEQ ID NO 514; 130pp; English.
 XX
 CC This invention, in the area of plant biotechnology, relates to novel
 CC polynucleotides comprising a nucleotide sequence encoding a protein which
 CC is involved in or associated with the synthesis, metabolism or
 CC degradation of carbohydrates in the plant grain and the expression of
 CC which is up-regulated during grain filling. The plant is selected from
 CC corn, tomato, banana, canola, cotton, peanut, sorghum, tobacco,

Db 186 LRAHNYKMPMMNDVRSLSALOGALRKABEHLGCLPADTPYSDFHHRFOELGLEKMG 245
Qy 242 DTAERVELMIOLLDLLEATDPCTLEKEFLGRIPMVFENVIIITPHGYFAQDNVLGYPDTGG 301
Db 246 DCARAGOSTLHLLDLLEAPDPSTLEKFLGTTIPMFENVIIISPHGYFAQDNVLGYPDTGG 305
Qy 302 QVVYILDDVRALENEMLRIRKQGGINTPRILITRLLIPDAVGTTCGGRLEKVGTEHSD 361
Db 306 QVVYILDDVRALENEMLRIRKQGGINTPRILITRLLIPDAVGTTCGGRLEKVGTEHSD 365
Qy 362 ILRVFPRKEGIVRWISRFKVPYLETYTEVDVAHEISKELHGPDLITGNXSDGNIVA 421
Db 366 ILRVFPRKEGIVRWISRFKVPYLETYTEVDVAHEISKELHGPDLITGNXSDGNIVA 424
Qy 422 SLAAKLGVTOCTTIAHALEKTYPPSDIYWKLEDKYHFSQCFADLFAMNHTDPIITST 481
Db 425 CLIAHKMGVTHCTIAHALEKTYPPSDIYWKLEDKYHFSQCFADLFAMNHTDPIITST 484
Qy 482 FOEINAGSKDTVGQYESHSTAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEEKRR 541
Db 485 FOEINAGSKDTVGQYESHSTAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEEKRR 544
Qy 542 LKHFRPELIDLLTYKVENEBHLCVLNDNRKPLFTMPRLDRVKNLTGLVEMCGKPKLRE 601
Db 545 LKSLHPELIEELLYSDVNDENKFLVLDKRNKPLIFSMARLDVRKNLTGLVEMCGKPKLRE 604
Qy 602 LANLVVGGDRRKESKDLBEKAEMKMFELIDKYNLNGQFPMWISSQMRIRNVELYRYIC 661
Db 605 LNVNLVVGCDHGNPDKDKEQHEQEPKMPDLIQYNLNGHIRWISQMRIRNVELYRYIC 664
Qy 662 DTGAFVOPALYEAFGLTVVEAMTCGLPTFATCNGSPAEIIVHGSGFNIDPYHGDQAD 721
Db 665 DTGAFVOPALYEAFGLTVVEAMTCGLPTFATCNGSPAEIIVHGSGFNIDPYHGDQAD 724
Qy 722 ILVDFEFCCKDPSHMDKISOGGLRIEKKYTWKLYSRELLLTGCVGPMKHSVLSERE 781
Db 725 LLVEFFECQGGHSHMTKISLGGIDRIEKKYTWKLYSRELLLTGCVGPMKHSVLSERE 784
Qy 782 SRRYLEMFYALKYRKLAESVPLAEE 806
Db 785 TRRYLEMLYALKYRKLAESVPLAEE 809

RESULT 12
AAE28500
ID AAE28500 standard; protein; 802 AA.
AC AAE28500;
XX
DT 27-DEC-2002 (first entry)
XX
DE Corn sucrose synthase shrunken-1 (Sh1) protein.
XX
KW Sucrose synthase; shrunken-1; Sh1; constitutive sucrose synthase; Sus1;
KW transgenic plant; plant breeding; grain breakage; grain strength; enzyme;
KW cellulose; corn; chromosome 9.
OS Zea mays.
XX
PN WO200267662-A1.
XX
PD 06-SEP-2002.
XX
PE 21-FEB-2002; 2002MO-US005137.
XX
PR 22-FEB-2001; 2001US-0270777P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Dhuga KS, Helentjaris TG, Niu X;
XX WPI; 2002-691625/74.
DR N-PSDB; AAD45851.

XX New polymucleotide and its encoded sucrose synthase, useful for
PT modulating the level of sucrose synthase in transgenic plants (e.g. maize
PT or soybean) to improve stalk length, reduce grain breakage, or improving
PT plant or grain strength.
XX
PS Example 9; Page 108-110; 125pp; English.
XX
CC The invention relates to two sucrose synthases, shrunken-1 (Sh1) and
CC constitutive sucrose synthase (Sus1) and their corresponding nucleic
CC acids. The polymucleotide, or its encoded protein, is useful for
CC modulating the level of sucrose synthase in a transgenic plant.
CC increasing cellulose production in the stalk tissue of a transgenic
CC plant, or increasing the concentration of cellulose in the tissues of a
CC seed of a transgenic plant. This is particularly useful in plants (e.g.
CC maize or soybean) breeding, especially for e.g. improving stalk length in
CC maize, reducing grain breakage during combining, transport or movement
CC into storage, or improving plant or grain strength. The present sequence
CC is corn Sh1 protein. Corn Sh1 gene is located at chromosome 9
XX
SQ Sequence 802 AA;
XX
Query Match 77.0%; Score 3268.5; DB 5; Length 802;
Best Local Similarity 76.2%; Pred. No. 2.4e-282;
Matches 609; Conservative 91; Mismatches 96; Indels 3; Gaps 2;
Qy 6 LTRVHSRERLDETLLAHRNELALSRIEKGKGIIOHQIILFEPAIPEENRKLKANG 65
Db 5 LTRHSRERLGAFFSSHPNELIALFSRYHQGGMLOHQHLLAEFDLPDSDEKKA-- 62
Qy 66 AFEEVLKASQBAIYLPWVALAVRPPQWMEYIRVNVYHALVBEIYTAEYHAFBEIYDG 125
Db 63 PFEDILPAQBAIYLPWVALAVRPPQWMEYIRVNVSELAIBEIYSEYLAFFKQILDG 122
Qy 126 SSNNFPLELDFEPPNSPFPPTLSKISGVVERLNLHSAKLPHDKSMKPLLEFRLVH 185
Db 123 QSNNSFPLELDFEPPNSPFPPTLSKISGVVERLNLHSAKLPHDKSMKPLLEFRLVH 182
Qy 186 CHKSKMMLNDRIQNLNALQHVLRKAEYLGTLPPETPCAFERHFOEIGLBERGNGDTAE 245
Db 183 NYKTTMMLNDRIQNLNALQHVLRKAEYLGTLPPETPCAFERHFOEIGLBERGNGDTAE 242
Qy 246 RYLEMIOLLDLLEATDPCTLEKEFLGRIPMVFENVIIITPHGYFAQDNVLGYPDTGGQVY 305
Db 243 RYLDLHLLDLLEAPDPSTLEKFLGTTIPMFENVIIISPHGYFAQDNVLGYPDTGGQVY 302
Qy 306 ILDOVRALLENEMLRIRKQGGINTPRILITRLLIPDAVGTTCGGRLEKVGTEHSDIRV 365
Db 303 ILDOVRALLENEMLRIRKQGGINTPRILITRLLIPDAVGTTCGGRLEKVGTEHSDIRV 362
Qy 366 PPRTEKGIWRKMSRFKVPYLETYTEVDVAHEISKELHGPDLITGNXSDGNIVASILA 425
Db 363 PPRNENGIRKMSRFPD-VMPYLETYTEVDVSEIMKEMQAKPDLITGNXSDGNIVATULA 421
Qy 426 HKLGVTCTIAHALEKTYPPSDIYWKLEDKYHFSQCFADLFAMNHTDPIITSTPEI 485
Db 422 HKLGVTCTIAHALEKTYPPSDIYWKLEDKYHFSQCFADLFAMNHTDPIITSTPEI 481
Qy 486 AGSKDTVGQYESHSTAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEEKRLGHF 545
Db 482 AGSKDTVGQYESHSTAFTLPGLYRVVHGIDVDFPKFNIVSPGADMEIYFPYTEEKRLGHF 541
Qy 546 HPEIEDLLTYKVENEBHLCVLNDNRKPLFTMPRLDRVKNLTGLVEMCGKPKLRELANL 605
Db 542 HPEIEBELIYSDVNSEHSEKFLVLDKRNKPLIFSMARLDVRKNLTGLVEMCGKPKLRELANL 601
Qy 606 VVVGDRRKESKDLBEKAEMKMFELIDKYNLNGQFPMWISSQMRIRNVELYRYICDTKG 665
Db 602 VIVAGDHGKESKDEEQHEQHEQEPKMPDLIQYNLNGHIRWISQMRIRNVELYRYICDTKG 661
Qy 666 AFVOPALYEAFGLTVVEAMTCGLPTFATCNGSPAEIIVHGSGFNIDPYHGDQADILVD 725
Db 662 AFVOPALYEAFGLTVVEAMTCGLPTFATCNGSPAEIIVHGSGFNIDPYHGDQADILVD 721

PF 01-MAR-1999; 99JP-00052102.
 XX
 PR 01-MAR-1999; 99JP-00052102.
 XX
 PA (MITA) MITSUI CHEM INC.
 XX
 DR WPI; 2000-675173/66.
 XX N-PSDB; AAC66090.
 PT Novel method for the reinforcement of cold resistance in a plant
 PT comprising introducing a vector encoding an enzyme that decomposes
 PT sucrose into the plant.
 XX
 PS Disclosure; Page 8-11; 22pp; Japanese.
 XX
 CC This invention relates to a method for the reinforcement of cold
 CC resistance in a plant, comprising introducing an expression vector having
 CC a DNA encoding an enzyme for decomposing sucrose connected downstream to
 CC a promoter expressible in the plant, and expressing the enzyme in the
 CC plant body. Included in the invention are an expression vector used in
 CC the method; a transformed plant carrying the expression vector; and a
 CC transformed rice carrying the above expression vector. The method is used
 CC for reinforcing cold resistance in a plant. The present sequence
 CC represents a sucrose synthase protein used in the method
 XX
 SQ Sequence 808 AA;

Query Match 76.7%; Score 3259.5; DB 3; Length 808;

Best Local Similarity 76.2%; Pred. No. 1.6e-281; Matches 609; Conservative 91; Mismatches 96; Indels 3; Gaps 2;

6 LTRVSLRERLDELTLAHRNELIALSLRLEGKGILOHIOIILEPEALPEENRKLKANG 65
 5 LARLSLRLERLQATSPSHNELIALFSRYVNGKGMLOHQLAEFDALREADKEKVA-- 62
 66 AFPEVLKASQEAIVPWPVALAVRPPGWEIYRNVVHALVVEILTVAEYLHFKELVDG 125
 63 PREDILRAQEAIVPWPVALAIRPPGWDYIRVAVSELAVESLSEYLAFLKQLVDG 122
 126 SNGNFVLELDEPPNPSFPPRTLSKISGNGVEPLNRHLSAKLFDKESMPLEFLRVH 185
 123 HTNSFPVLELDEPPNPSFPPRPSKISGNGVQFLNRHLSKLPDCKSLYILFLKQH 182
 186 CHKGKMMMLNDRIQMLNLQHLVRAKAEYLGTLPEPTCAFEHREFOEIGERNGD7AE 245
 183 NHKGTMMMLNDRIQSLRGLQSLRAKAEYLGIMIPDTPYSEHNRFOELGEEKGMDCAK 242
 246 RYLEMIQILLLELLEATDPTLEKPIGRIPWVNVVILTPHGFADNDVLGYDTGGQVYV 305
 243 RVLDTIHLLELLEAPDPANLEKPIGTLPMFNVVILSPHGFADSNVLYGPDGTGGQVYV 302
 306 IIDQVRALENEMMLRIKQGLNITPRILITLRLPDVAGTGGORLEKYGTEHSDILRV 365
 303 IIDQVRALENEMMLRIKQGLDITPKILITVLLPDVAGTGGORLEKYGTEHSDILRV 362
 366 PERTEKGIYKWMISREKVPYLETYETEDVAHEISKEILHGTBDLIIGNXSDGNIVASLLA 425
 363 PERSENGILRKWISREPD-VMPLETYETEDVAHEIREMOKAKDILIIGNVSDGNIVATLLA 421
 426 HGLSTVQCTIAALAEKTKYPPSDIYWKLEDKYHNSCOPTALFPMNHTDPIITSTFOEI 485
 422 HGLSTVQCTIAALAEKTKYPPNSDIYLDKFDSDYHNSCOPTALIMNHTDPIITSTFOEI 481
 486 AGSKDTGVGESHHTAFTLPGLRVAVHGDVDPKKNIVSPGADMEIYPPYTEKSKRLKHF 545
 482 AGSKDTGVGESHHTAFTLPGLRVAVHGDVDPKKNIVSPGADMEIYPPYTTADKRLAF 541
 546 HPEIEDLLYTKVNEEHLCLVNDKRPILFTMPRLDRVNLTLGLVESCCKNKLRELANL 605
 542 HPEIEELLYSEVENEDEHFKVLKDKNKPITFSMARLDRVNMGLVEMYGKNAHLADLANL 601
 606 VVVGDRRRESKDLBEKAKMKKPFELIDKYNLNGQPRWISQGMIRVVELYRYICDTRG 665

DB 602 VIVCGDHNGSKREDEQAEFKOMGLIDQYKHKHIMWISQNMNRVANGELYRYICDTRG 661
 QY 666 AFVOPALYEAFTLVVEMTGLPTFATCNGSPAEIIVHGKSGFNIDPYHSDOADIIVD 725
 DB 662 VFQAPAFYEAFTLVVEMTGLPTFATCNGSPAEIIVHGKSGFNIDPYHSDOADIIVD 721
 QY 726 FFEKCKDPSHWDKISQCGKRIEKKYTWKYSRLLTLTSVYGFWMKVSLSLERESARY 785
 DB 722 FFEKCKDPSHWDKISQCGKRIEKKYTWKYSRLLTLTSVYGFWMKVSLSLERESARY 781
 QY 786 LEMFYALKYRKLAEVPLA 804
 DB 782 IEMFYALKYRKLAEVPLA 800

RESULT 15

ADCC68459

ID ADCC68459 standard; protein; 808 AA.

AC ADCC68459;

DT 18-DEC-2003 (first entry)

XX *Lolium perenne* fructan biosynthesis protein SEQ ID NO:169.

XX *Lolium perenne*; *Festuca arundinacea*; *Lignin*; *fructan*; *tannin*;

XX *biosynthetic pathway*; *plant*.

XX *Lolium perenne*.

XX 15-MAY-2003.

XX 07-NOV-2002; 2002MO-NZ000239.

XX 07-NOV-2001; 2001US-0337703P.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX (WRIG-) WRIGHTSON SEEDS LTD.

XX Demmer J, Forster RL, Gibson JB, Shenk MA, Norris MG, Glenn M,

XX Sautsbury KM, Hall C;

XX WPI; 2003-441544/41.

XX N-PSDB; ADCC68423.

XX New polynucleotide encoding polypeptides from *Lolium perenne* or *Festuca*

XX *arundinacea*, useful for modulating the biosynthesis of *lignin*, *fructan* or

XX *tannin* in a plant.

XX Claim 18; SEQ ID NO 169; 240pp; English.

XX The present invention describes isolated polynucleotides (I) encoding

XX proteins (II) from *Lolium perenne* and *Festuca arundinacea* which are

XX active in *lignin*, *fructan* and *tannin* biosynthetic pathways. Also

XX described: (1) an isolated oligonucleotide probe or primer comprising at

XX least 10 contiguous residues complementary to 10 contiguous residues of

XX (1); (2) a kit comprising the oligonucleotide probe or primer; (3) a

XX genetic construct comprising (1); (4) a transgenic plant cell comprising

XX the genetic construct of (3); (5) a plant or its seed, fruit or progeny

XX comprising the transgenic plant cell of (4); (6) modulating one or more

XX of the *lignin*, *fructan* or *tannin* compositions of a plant; (7) producing a

XX plant having one or more of the *lignin*, *fructan* or *tannin* compositions;

XX and (8) modifying the activity of (ii) involved in a *lignin*, *fructan* or

XX *tannin* biosynthetic pathway in a plant. (I) can be used for modulating

XX the biosynthesis of *lignin*, *fructan* or *tannin* in a plant. The present

XX sequence is used in the exemplification of the present invention.

XX Sequence 808 AA;

XX Query Match 76.6%; Score 3254.5; DB 7; Length 808;

XX Best Local Similarity 76.2%; Pred. No. 4.4e-281;

Matches 609; Conservative 92; Mismatches 95; Indels 3; Gaps 2;

QY	6	LTRVSLBERLDEJTLAHHNELITALSTRLEGKGILOHNOHIIIEEALPEENKRYL	65
Db	5	LTRLHSLERLCATSSHNELIALSKYVHOGKQLOHOLTEELPEADKERYA--	62
QY	66	AFVEVLKASOEAVILPRWVALAVRPRPGWMEYIRVNVHVALVBEELTVAEYHKEELVDG	125
Db	63	PEEDILBRAQGEAVILPRWVALAIRPRPGWMDYIRVNVSELAVELTVSEYLAKEQOLVDE	122
QY	126	SSNGNFVLELDEPPFNSSFPRLPSKISNGCVFELNRHISAKLPHDKESMHLPLEBRYH	185
Db	123	HASSKFVLELDEPPFNASFPRLPSMSKISNGVOFLNRHSKFLPDKESYPLNLELKAH	182
QY	186	CHKKNMMLNDRIONUNALOYHVRKKEEYLGTLPRETPCAEFHRFOEIGLERGMGDTAE	245
Db	183	NHKGTMLNDRIOISLRGQSALRKEEYLTSLPEBTPSSEFHRFOELGLEKMGEDTAK	242
QY	246	RYLEMIOLLLDLLEATDPCTLEKFLGRIPMFVNVILTPHGYAODNVLGYPDTGQVUY	305
Db	243	RVODTILHLLDLLEADDPASLEKFLGTIPMFVNVILSPHGYAQSNVLGYPDTGQVUY	302
QY	306	ILDOVALENEMILRKIOOGLNITPRILITRLPRAVGTGQORLEKYGTSHSLRY	355
Db	303	ILDORALENEMILRKIOOGLDITPRILITRLPRAVGTGQORLEKYGTSHSLRY	352
QY	366	PRTEKGIARKKISIREKWPYLETETEDVAHEISKEHGTPOLIIGNSDGNIVASLLA	425
Db	363	PRTEKGIARKKISIRD-VMPYLETETEDVAHEINEMQOTKPRLIIGNSDGNLVAITLA	421
QY	426	HKLGYTOCTIAALEKTKYPSDSIYWKKLBDKYKHSOCTADL.PAMNHTDPIITSTPOET	485
Db	422	HKLGYTOCTIAALEKTKYPSNSDIYADKEDPSQYHSCOFADLIAMNHTDPIITSTPOET	481
QY	486	AGSKDTVGYESHTAFTLPGLYRVVHGIDVPDKFNIVSGADAMELYPPTEEKREBLKHF	545
Db	482	AGSKDSVGOYESHIAFTLPDLRYVHGIDVFDKFNIVSBGADMTYVFFTEEDKLLTAF	541
QY	546	HEIEDLLTYKYNEBEHLCVLDRNRPILFTMPRLDRVYNLTGLVEMCGNPKYRELANTL	605
Db	542	HPEIEELLYSDVENSEBHKFVLKDKNRPILFSMARLDRVKNMTGLVEMFGNAHLKOLANTL	601
QY	606	VVVGDDRRESKDLBEKAMKMKFELIDKYNLNGORWATISSOMNRRIRNVELYXICDTKG	655
Db	602	VIVADDHGESKDREREQAEFFKNYSILIEEKLEGHIRWLSAOMNRRARNELXYICDTKG	651
QY	666	AFVQALYEAFFGLTVVEAMTCGLPTPATNGGGAELIIVHGKSGSFNIDPYHGDAADILVD	725
Db	662	AFVQAFYEAFFGLTVVEAMTCGLPTATCHGGAELIIVNGVSGSLAHDIPHSDQAAOILVN	721
QY	726	FEKCKKDPBHWDKISQGLKRIIEKYTWKIVSERLLTLTGYYGFWKHAVSNLERRSRRY	785
Db	722	FFEKSTADPTYMDKMSBEGLKRIYEKYTWKIVSERIMLTITGVYGFWKVSNLERRSRRY	781
QY	786	LEMFAKTKRKLAESVPLA 804	
Db	782	LEMFAKTKRSLAAVPLA 800	

Search completed: June 2, 2004, 14:51:05
Job time : 66 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 14:49:19 ; Search time 21 Seconds
(without alignments)
3691.922 Million cell updates/sec

Title: US-10-003-405-2
Perfect score: 4247
Sequence: 1 MAERALTQVHSLRERLDETL.....EMFYALKYRKLAESVPLAEE 806

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3669.5	86.4	806	S31479	sucrose synthase (
2	3614.5	85.1	808	T49233	sucrose synthase-1
3	3501.5	82.4	805	YUPOS	sucrose synthase (
4	3400.5	80.1	808	S37560	sucrose synthase (
5	3268.5	77.0	802	YUZMS	sucrose synthase (
6	3264	76.9	816	S31939	sucrose synthase (
7	3259.5	76.7	808	S23543	sucrose synthase (
8	3238.5	76.3	801	T14338	sucrose synthase (
9	3222	75.9	816	S32451	sucrose synthase (
10	3209	75.6	807	S29242	sucrose synthase (
11	2985	70.3	809	T06497	sucrose synthase (
12	2962	69.7	809	B85029	probable sucrose s
13	2858	68.6	822	S71493	sucrose synthase (
14	2892	68.1	808	T01420	sucrose synthase (
15	2815	66.3	804	YUWU	sucrose synthase (
16	2479.5	58.4	942	S24966	probable sucrose s
17	2312.5	54.5	986	C96760	sucrose synthase (
18	1747	41.1	806	A12428	sucrose synthase (
19	1435.5	33.8	805	AH1938	sucrose synthase (
20	941	22.2	235	JT0281	sucrose synthase (
21	883	20.8	212	JT0280	sucrose synthase (
22	803.5	18.9	204	S22535	sucrose synthase (
23	756	17.8	179	S22537	sucrose synthase (
24	661	15.6	218	S22131	sucrose synthase (
25	621	14.6	127	A29484	sucrose synthase (
26	494.5	11.6	720	S75935	hypothetical prote
27	437	10.3	1059	T12195	sucrose-phosphate
28	436	10.3	1068	J01329	sucrose-phosphate
29	435.5	10.3	1083	T04062	sucrose-phosphate

30	431	10.1	1056	2	J02277	sucrose-phosphate
31	421.5	9.9	1045	2	S55253	sucrose-phosphate
32	418.5	9.9	1081	2	T09837	sucrose-phosphate
33	413.5	9.7	1053	2	S34172	sucrose-phosphate
34	407	9.6	1057	2	S72648	sucrose-phosphate
35	404	9.5	1047	2	T51800	sucrose-phosphate
36	400	9.4	1084	2	T04103	sucrose-phosphate
37	394	9.3	1054	2	T09833	sucrose-phosphate
38	387.5	9.1	1049	2	J04783	sucrose-phosphate
39	386	9.1	1064	2	B6182	hypothetical prote
40	294	6.9	100	2	T09856	sucrose synthase (
41	256	6.0	67	2	T12251	sucrose synthase (
42	248	5.8	454	2	AG2657	glycosyltransferase
43	248	5.8	454	2	D97439	hypothetical prote
44	206.5	4.9	501	2	T01981	sucrose-phosphate
45	197.5	4.7	341	2	S72649	sucrose-phosphate

ALIGNMENTS

RESULT 1						
S31479						
sucrose synthase (EC 2.4.1.13) - fava bean						
C.Species: Vicia faba (fava bean)						
C.Date: 02-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 18-Jun-1999						
C.Accession: S31479						
R.Heim, U.; Weber, H.; Wobus, U.						
Submitted to the EMBL Data Library, December 1992						
A.Description: Sucrose synthase expression patterns in developing seeds of Vicia faba in						
A.Reference number: S31479						
A.Accession: S31479						
A.Status: preliminary						
A.Molecule type: mRNA						
A.Residues: 1-806 <HEI>						
A.Cross-references: EMBL:X69773; NID:G22037; PIDN:CAA49428.1; PID:G22038						
C.Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology						
C.Keywords: glycosyltransferase; hexosyltransferase						
F.273-753/Domain: sucrose-phosphate synthase homology <SSFS>						
Query Match						
Best Local Similarity 86.4%; Score 3669.5; DB 2; Length 806;						
Matches 695; Conservative 52; Mismatches 58; Indels 1; Gaps 1;						
QY	1	MAERALTQVHSLRERLDETLAHRNEIALLSRIBGKGGIIOHQIILFEPAIPEERK	60			
DB	1	MATERLTQVHSLRERLDETLTANNEIALLSRIBGKGGIIOHQIILFEPAIPEERK	60			
QY	61	KLANGAPEVTKASQEAIVLPWVALAVRPGWERYRNVVHALVVEELTVAEYLHPRKE	120			
DB	61	KLTDGAFGEVRSSTQEAIVLPWVALAVRPGWERYRNVVHALVVEELTVAEYLHPRKE	120			
QY	121	ELVGGSSNGNFVLBIIDPEFPNSSFPRPTLSKISGVGFELNRHLSAKLFHDKESMPLLE	180			
DB	121	ELVGGSSNGNFVLBIIDPEFPNSSFPRPTLSKISGVGFELNRHLSAKLFHDKESMPLLE	180			
QY	181	FLRVHCHGKMMMLNDRIQNLALOHVLRKABEYIGTLPPETPCAFEFHRRQIELENGW	240			
DB	181	FLRVHCHGKMMMLNDRIQNLALOHVLRKABEYIGTLPPETPCAFEFHRRQIELENGW	240			
QY	241	GDTARVLEMTQLLDLEATDPCTLEKFLRIPVNFVNLITPHGYFAQONNVGYPTDG	300			
DB	241	GDSARVLEMTQLLDLEATDPCTLEKFLRIPVNFVNLITPHGYFAQONNVGYPTDG	300			
QY	301	GQVAVIILDQVRLALENEMLRIRKQOGLNTPRILITRLLPDAVGTTCGRLEKYGTSHS	360			
DB	301	GQVAVIILDQVRLALENEMLRIRKQOGLNTPRILITRLLPDAVGTTCGRLEKYGTSHS	360			
QY	361	DILRVPTTEKGIYRKATSRPEKWPYLETTEDEVNAHISKEMLGTPRLIIGNSDGIIV	420			
DB	361	DILRVPTTEKGIYRKATSRPEKWPYLETTEDEVNAHISKEMLGTPRLIIGNSDGIIV	420			
QY	421	ASLAAKLGVTGCTAAHLEKTKYRPPDSIDYWKKLEDKXHFSCQFTADLFAMNHTDFTTS	480			

Db	420	ASLAAHKLGVQCTTAHLEKTKYPSDIIYAKKEEEKHSCQSTADI.PAANHHDPIITTS	479
Qy	481	TFQEIAGSKDVGQYESHATFTLLPGLYRVHGVIVDFPKENIVSPGADMEIYFPYTEKR	540
Db	480	TFQEIAGSKDVGQYESHATFTLLPGLYRVHGVIVDFPKENIVSPGADQTIYFPYTELSR	539
Qy	541	RLKHFHPEIEDLLTYKVENEBELCVLANDRNKPIILFTMRDLRVNKL.TGLVEMCGKNPKLR	600
Db	540	RLTSPFYPIREIELLYSTVENEBHICVLKDRSKPIIFTMARLDRVKNITGLVEMYGKNKLR	599
Qy	601	ELANL.VVVGDDRRKESKDLBEKAEMKKMFELIDKINLNGQFRWISSQMRIRNVELYRI	660
Db	600	ELVNLVVVAVGDRRKESKDLBEKAEMKKMYELIELTYLNGQFRWISSQMRIRVNGELYRI	659
Qy	661	CDTYGAFVQPLLYEAFGLTVVEAMTCGLPTPATCNGCPAEIIVHGKSSFNIDPHGDOAA	720
Db	660	CDTGAFVQPAVVEAFGLTVVEAMATGTPPTATLNGCPAEIIVHGKSGFHIIDPHYGRDAA	719
Qy	721	DILVDFEKKCKDDSHMDKISIQGGIKRIIEKTYMKIYSEBLITLTGVYGFPMKHVSNLERR	780
Db	720	DILVEFFEFKVADSHMDKISIQGGIQRIIEKTYMKIYSORLITLTGVYGFPMKHVSNLDRL	779
Qy	781	ESRRYLEMFYALKYRKLAESVPLAE 806	
Db	780	ESRRYLEMFYALKYRKLAESVPLAVE 805	

RESULT 2

149233
sucrose synthase-like protein - *Arabidopsis thaliana*
N/Alternate names: protein F7KL5.40
N/Species: *Arabidopsis thaliana* (mouse-ear cress)
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jun-2000
C/Accession: T49233
R/Organism: *B. oleracea*, *B. oleracea*, D.; Zeidler, K.; Mewes, H.W.; Rudd, S.; Lem
submitted to the Protein Sequence Database, April 2000
A/Reference number: 225019
A/Accession: T49233
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-808 <OBS>
A/Cross-references: EMBL:AL353871; GSPDB:GN00061; ATSP:F7KL5.40
A/Experimental source: cultivar Columbia; BAC clone F7KL5
C/Genetic8:
A/Gene: ATSP:F7KL5.40
A/Map position: 3
A/Intons: 357/2, 128/2, 192/3, 304/3, 336/3, 394/3, 433/3, 489/2, 564/2, 752/2, 798/3
A/Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology

Query Match	85.1%	Score 3614.5	DB 2	Length 808
Best Local Similarity	84.6%	Pred. No. 1.2e-211		
Matches 681	Conservative 58	Mismatches 65	Indels 1	Gaps 1
QY	2	AERALTTRVHSDRELRDETLTLLAHNEITALLSRIGEGKGIILOHHQIIIEFEALPEENRKK	61	
DB	4	AERVITTRVHSGEREDATITVAQKNEVPALLSRVEAKGIILOHHQIIIEFEAMPLETQKK	63	
QY	62	LANGAFPEVLKXSQSAIYLPWPVALAVRPRGVWYIIVNVHVALVVEALTVAEYTHPEE	121	
DB	64	LKGGFFPEFLNSAQAIYLPWPVALAVRPRGVWYIVRVNHLADVLEELQNSYIQFEE	123	
QY	122	LVDSGNSGNFVLELDPEEPNSSFPRPTLSKISGNGVEFLNRHLAKLPHDKESMHPLEEF	181	
DB	124	LVDGIKNGNFTLELDFEPPNNAFPRPTLNTKYIGDVEFLNRLSKLPHDKESLHPLEKF	183	
QY	182	LRVGHKKKMMNLNRIQNLNLALCHVLRKAEYITGLTPERTPCAEFRHPOEIGERGMG	241	
DB	184	LRVHHEGTEGLMLNRIQNLNLALCHVLRKAEYITGLTPERTLYSFEHKKFOEIGERGMG	243	
QY	242	DTAEVLEMIQLLDLLEATDPCTEKFLGRIPWPNVVIITPHSGFAQDNYLYGPDTCG	301	
DB	244	DTAEVLEMLNRIQLLDLLEAPDCTLENFGRIIPWPNVVIITSPHSGFAQDNYLYGPDTCG	303	

QY	302	QVWYILDDOVRALENEMLRIRKQOGNITPRILIIITRLRLPDVAGTTGGOLERKYGTGEHSD	361
Db	304	QVWYILDDOVRALETEMLRIRKQOGNITPRILIIITRLRLPDAGATTGGOLERKYGSQYCD	363
QY	362	ILRVFPRTEKGIVRKMIRFEKWPVYLETVEDVAHEISKLAGPTDILIIGNXSDSNIVA	421
Db	364	ILRVFPRTEKGIVRKMIRFE - VWPYLETFTEDVAHEISKLEQKPDILIIGNXSDGNLVA	422
QY	422	SILAHKLVCTOCTIIHALEKTKYPPSDIYWKLEDBKYNHSCOFIADLIFMNHDTFIITST	481
Db	423	SILAHKLVCTOCTIIHALEKTKYPPSDIYWKLEDBKYNHSCQFTADLIIAMNHDTFIITST	482
QY	482	FOEINGSKDVOGYEYESHATFTLLPGILRVVHGIDVPDPKENIVSPGADMEIYEPYTEEKR	541
Db	483	FOEINGSKDVOGYEYESHRSFTLLPGILRVVHGIDVPDPKENIVSPGADMSIYFAYTEEKR	542
QY	542	LKHHPREIBEDLIYTYENBEHLCVINDRNKPLIFTMPLRLDRKYNLTGLYEWCKSNKRLRE	601
Db	543	LTAFLHEIBELLYSVENEBEHLCVAKDKKKPLIIFTMARLDRVYNLSGLVEWYGKNTRLRE	602
QY	602	LANIVVGGDRPKESKDLLEKAKEMKMPFLIDKYNLNGOPRMTSSOMNRIINAVELRYIC	661
Db	603	LVNLVVVGGDRPKESQDNEEKAKEMKMPHELIEEYKLNNGOPRMTSSOMNRRINCELRYIC	662
QY	662	DTKGAFFVOPALYEAFGLTVEAFMTCGLPTFAFCNGGPABIIYHGKSGFNIDIRPHGOAAD	721
Db	663	DTKGAFFVOPALYEAAGLTVEAMTCGLPTFAFCNGGPABIIYHGKSGFNIDIRPHDKAAE	722
QY	722	ILVDFEKKCKDPSHMDKISOGGLKRIIBEKYTWKIYSEBLITLTVGYGFWKHSVNLERRE	781
Db	723	SLADPFITCKKHDPSHMDQISLGGLEIRIOGKTYMIOISGRLITLTVGYGFWKHSVNSLDRLE	782
QY	782	SRRYLEMPALKYRKLAESVPLAE 806	
Db	783	SRRYLEMPALKYRPLAQAVPLAE 807	

RESULT 3

sucrose synthase (EC 2.4.1.13) - potato
 C.Species: Solanum tuberosum (potato)
 C.Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 11-Jun-1999
 C.Accession: A29615
 R.Salanoubat, M.; Belliard, G.
 Gene 60, 47-56, 1987
 A.Title: Molecular cloning and sequencing of sucrose synthase cDNA from potato (Solanum
 A.Reference number: A29615; MUID:88152501; PMID:2964386
 A.Accession: A29615
 A.Molecule type: mRNA
 A.Residues: 1-805 <SAL>
 A.Cross-references: GB:M18745; NID:G169571; PID:AAA3841.1; PID:G169572
 A.Experimental source: var. Sirtema 2n=4x
 C.Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
 C.Keywords: glycosyltransferase; hexosyltransferase
 F.1279-753/Domani: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match	82.4%;	Score 3501.5;	DB 1;	Length 805;
Best Local Similarity	82.1%;	Pred. 3,8e-224;	Mismatches 78;	Indels 1;
Matches	662;	Conservative	65;	Matches 24;
Qy	1	MAERLALTRHSLRERLDFTLLAHNRNELLALSLRISBEKGKGLIOHQHIIIEFEALPIEENRK	60	
Db	1	MERVLTRHSLRERVDATLLAHNRNELLFLSRISBEHGKGLKPHLELAFDRIRODDN	60	
Qy	61	KLANGAFEEVLKASQEAIVLPVWVALAVAPRPQWMEYIRVNVALVVEELTVAEYIAHFKS	120	
Db	61	KLNRHAFEEELASTQGAIVLPVWVALAIIILRPQWMEYIRVNVALVVEELSVPEYLQFKE	120	
Qy	121	ELVDGSSNGNFVLELDPEFPNNSFPPEPTLSKISGNGVEFLNRHLAKLFLDKSSMHPLE	180	
Db	121	ELVDGSSNGNFVLELDPEFPFTASFPKPTLLTKISGNGVEFLNRHLAKAMFDDKSMTPLE	180	

A1:Map position: 9
A1:introns: 32/2; 72/3; 123/2; 187/3; 227/2; 299/3; 331/3; 389/3; 428/3; 484/2; 559/2; 661/2
C1:Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
C1:Keywords: glycosyltransferase; hexosyltransferase
F1:276-750/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match	77.0%;	Score 3268.5;	DB 1;	Length 802;
Best Local Similarity	76.2%;	Pred. No. 1.1e-208;		
Matches 609;	Conservative 91;	Mismatches 96;	Indels 3;	Gaps 2

QY	6	TRVHSLEBRDETLHAHNEIALLSRIIEGGKIILOHOOIIIEFEAIIPRENNKRLANG	65
Db	5	LTRLHSLEBRICATISSHPNEILALFSRYVHOGKMLORHOLIAEFDLFSDXEKTA--	62
QY	66	AFFEVYLKASOEAIYLP.PWVALAVRDPGVMEYIRVUNVALVEELTVAEYLHFKKEILDVG	125
Db	63	PFEDILRAAQEAIYLP.PWVALAIRDRPGWMDIYIRVNSVELAVEELSVEYLAFKEQLVDG	122
QY	126	SSNGNFVLELDPEPNSSFPFRPTLSKINGVEFLNRHLSAKLPHDKSSMHPLLFELRVH	185
Db	123	QSNSNFVLELDFEPNNAFPRPSMSKSIINGNQFNRHLSKLPDKSKSLPLPLNLFVAH	182
QY	186	CHKGNMMLNDRIOULNALOHVLRKAEYELGILPETPCAERENHFOEIGYERGMDTAE	245
Db	183	NYKGTMMMLNDRIOGLRGLQSSLRKAEYELSVDPDITYSEFNHRFOELGLEKNGDTRAK	242
QY	246	RVLEMIOILLDLLEATDPCLTLEKFLGRI.PWENVVILTRPHGYAODNVLGY.PDTGGQVY	305
Db	243	RVLDITLHLLDLLEAPDRPANLEKFLGTIPMMENVVILSPHGFAQSNVLGY.PDTGGQVY	302
QY	306	ILDOVRALENEMLLARIKQOGLNITPRILILTRLLPDAAGTCQGLLEVYGTENHSDILRV	365
Db	303	ILDQVRALENEMLLARIKQOGLDITPKILIVRLLPDAAGTCQGLLEVYGTENHSDILRV	362
QY	366	PFRTEKGIYVRKWIISFEKAWPYLETYTEDDVAHEISKELAHGT.PDLIIGNXSDGNIVASLLA	425
Db	363	PFRNNGILRKKIISFDP--VWPYLETYTEDVSEIIMKENQAKPDLIIGNYSOGNIVATLLA	421
QY	426	HKLGYTOCTIAHALEKTRY.PDSDIYWKLEDKTKHFS.COPTADLFAMNHTDPIITSTFOEI	485
Db	422	HKLGYTOCTIAHALEKTRY.PNSDIYLDKFPDSGYHFS.COFTADLLAMNHTDPIITSTFOEI	481
QY	486	AGSKOTVOYESHTFTL.PGLYRVVHGI.DVPRKKNIVYSPGADMEITYPYTBEEKRLHGF	545
Db	482	AGSKOTVOQYESHILFTL.PGLYRVVHGI.DVPRKKNIVYSPGADMSVYPIYETDTRKRLAF	541
QY	546	HPEIEDLLYTKVENEHLCVILDRNKPIL.FTWPRLDRYKNLTGLVEMOGKPKJRELANTL	605
Db	542	HPEIIEILYSDVENSEHKFVILKDKKKKPIIFSMARLDRYKNMTGLVEMYGXNARLRELANL	601
QY	606	VYVGGDRKRESKOLLEKBAKMKMFELIDKYNLNGCFNMISSQMRIRANVELYRYICDTKG	665
Db	602	VIVADHDHESKDXREEOAEFKKMYGLIDERYKLKGIRWISAKQMNVRNGELERYRYICDTKG	661
QY	666	AFVQALYEA.FGLYVEAMTGL.PTFATCNGGPAIIYHAGSGFNIDPYHGQADAILVD	725
Db	662	AFVQAFYEA.FGLYIESMTGSLPTIATCHGSPAILIYDGVSGHLIIDPYHSDKAADILVN	721
QY	726	FEFKCKDPSHWDKISQGGKLRIEEKYTWKIYSEBLLTLTGYYGFWKXVSNLERRSRY	785
Db	722	FFDCKADPSYWDIEISQGGKQRIYEEKYTWKIYSEBLLMTLTGVYGFWMKYVSNLERRTRY	781
QY	786	LEM.FALYKRLKLAESVPLA 804	
Db	782	IE.MFALYKRLASQVPLS 800	

RESULT 6
S19139
sucrose synthase (EC 2.4.1.13) 2 - rice
N:Alternate names: sucrose-UDP glucosyltransferase 2
C:Species: Oryza sativa (rice)
C:Date: 19-Mar-1997 #sequence, revision 18-Jul-1997 #text, change 18-Jun-1999

C:Accession:SI19139; S77947; S77943; S22536
R:Yu, W.P.; Wang, A.Y.; Jiang, R.H.; Sung, H.Y.; Su, J.C.
Plant Mol. Biol. 18, 139-142, 1992
A:Title: Isolation and sequences of rice sucrose synthase cDNA and genomic DNA.
A:Reference number: SI19139; MUID:92119223; PMID:1531032
A:Accession: SI19139

A: Molecule type: DNA
A: Residues: 1-816 <YUL>
A: Cross-references: EMBL:X59046
A: Accession: S77947
A: Molecule type: mRNA
A: Residues: 1-816 <YUL>
A: Cross-references: EMBL:X59046
R: Su, J.C.

submitted to the EMBL data library, April 1991
A:Reference number: S77943
A:Accession: S77943
A:Molecule type: DNA
A:Residues: 1-447 'Y', 449-816 <SNU>
A:Cisotrans-references: EMBL:X59046; NITD:g20094; PIDN:CAA41774.1; PID:g20095
R:Wang, A.-Y.; Yu, W.-P.; Chung, R.H.; Huang, J.W.; Sung, H.-Y.; Su, J.-C.

Plant Mol. Biol. 18, 1191-1194, 1992
A1>Title: Presence of three rice sucrose synthase genes as revealed by cloning and sequencing
A1.Reference number: S22535; MUID:92288314; PMID:1534703
A1.Accession: S22536

A;Status not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 605-816 <WAN>
A;Experimental source: cv. Tainong 67

C:Genetics: 37/2; 80/3; 131/2; 195/3; 235/2; 307/3; 339/3; 397/3; 436/3; 492/2; 567/2; 67
A:Interome: sucrose synthase; sucrose/sucrose-phosphate synthase homology
C:Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
C:Keywords: glycosyltransferase; hexosyltransferase
E:284-758,Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match	76.9%	Score 3264	DB 2	Length 816
Best Local Similarity	76.0%	Pred. No. 2.2e+08		
Matches 612; Conservative	91	Mismatches 100	Indels 2	Gaps 2

```

QY 3 ERALTRVSHLEBRLDETLTAHNRHETLALSRILEGGKGILOHNOIIEFF-EAIPENRKK 61
Dp 7 DVALSRILSVBRERIGDSLASHAPNEIIVAAFTVLVLGKMLQHOQIIAETVNNASIADEKK 66
QY 62 LANGAFPEVLVASOEAIPLPMAVLAVAPREGEVEYIRAVNHALVBEELTVAEYLHFFEE 121
Dp 67 LKDGAFPEVLVASOEGEIVISPMVALAIRRGVEVEYAVNNSELAVELLYTPEYLOFQEQ 126
QY 122 LYDSSNGNFWLELDFEFPNNSFPRLPSKIGNGVEFLNHLAKLFDHXSHPLLLEF 181
Dp 127 LVEEGTNNFWLELDFEFPNASFPRLPSKIGNGVQFLNHLHLSKLFHDHXSMPPLNF 186
QY 182 LRVHGHKKMMNDRIQNLALQHVLRKABEYLTLPREPLCAEFHRPOEIGERQWG 241
Dp 187 LRANHYKMTVMNDLRIRLSALQALKAEBHLSGLSADRPYSSEHHRFOLGIEKQWG 246
QY 242 DTAERVLEMIQILDLIELBATPCTLEKEJGRIPMFVFNVLITPHGYFADQVNLGYRDPGG 301
Dp 247 DCAKRSQETIHLIDLLEAPRPSLEKLGITPMVFNVNINSPHCYFQANVLGIPDGG 306
QY 302 QVVYLIDQVRALENEMLLIRIKOQGLNTPRILITRLLPDAVGTTCGRLEKVEYGTESHD 361
Dp 307 QVVYLIDQVRALENEMLLIRIKOQGLNTPRILITRLLPDAVGTTCGRLEKVEYGTETH 366
QY 362 ILVAFPRLEKGIYKRWISREKWPYLETTEYEDVAHEISKLHGPRDILIGNXSGNIVA 421

```

Db	367	ILIRVPERENGIYVRKMISRFE-VMPEYLTFFHDDVAHEIAGELIQANPDIIINISGNGLVA	425
QY	422	SLIAHKLVGTQCTAAHALEKTYRPPSDIYWKKLBDKYHPSQCFADLEFAMNHTDFIITST	481
QY	436	CLIAHKGVCTCTAAHALEKTSKPPSDIYMKKFBEDHYHPSQCFYTDLLAMNHADIITST	485
QY	482	FOELAGSKDTGVQYESHTAFTLPGIYRVVHGIDVDPKENTVSPGADNIEIYFPYTEKRR	541

Db 486 FOEINAKDVTQGYQYSHNAFTWPGLVVHGIDVDPKFNIVSPGADMSIYEPYSESRRK 545
 Qy 542 LKHPRPTEIDLLTYTVENEHLCVANDRNKPLFTMPPLDRYKNTGLVWNGKPKARE 601
 Db 546 LKSHPELEELLYSSVDNNHFKMLKDRNKPFIIFSMARLDRYKNTGLVWNGKPKARE 605
 Qy 602 LANLVVVGDDRRKESKDLSEKAKMKMFELIDKYNLMGOFMWISSQMNRIKRVNLYRYIC 661
 Db 606 LANLVVVGDDHGNPKDKEQAEQKFMFDLIHQYNLNGHIRKISQMNRIKRVNLYRYIC 665
 Qy 662 DTGAFFVAPALYEAAGLTIVBEAMTGLPTFATCNGGPAEIIYHGKSGFNIDPYHQDQAD 721
 Db 666 DTGAFFVAPALYEAAGLTIVBEAMTGLPTFATCNGGPAEIIYHGKSGFNIDPYHQDQAD 725
 Qy 722 ILVDFEKKCDPSHMDKISQGLKRIEKKYTWKLYSERLLTLTGVGFWKVSUMLERRE 781
 Db 726 ILVDFEKKCDPSHMDKISQGLKRIEKKYTWKLYSERLLTLTGVGFWKVSUMLERRE 785
 Qy 782 SRRYLFMYALKYRKLAESVPLAE 806
 Db 786 TRRYLEMLYALKYRKLAESVPLAE 810

RESULT 7

S23543
 sucrose synthase (EC 2.4.1.13) 1 - rice
 NAlternate names: sucrose-UDP glucosyltransferase 1
 CSpecies: Oryza sativa (rice)
 CDate: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
 CAccession: S23543; S25526
 R.Wang, M.B.; Boulter, D.; Gatehouse, J.A.
 Plant Mol. Biol. 19, 881-885, 1992
 A>Title: A complete sequence of the rice sucrose synthase-1 (RSS1) gene.
 A.Reference number: S23543; MUID:92153399; PMID:1386537
 A.Accession: S23543
 A>Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-808 <MAN>
 A.Cross-references: EMBL:X64770; NID:g20365; PIDN:CAA46017.1; PID:g20366
 R.Odegaard, W.; de Lumen, B.O.
 submitted to the EMBL Data Library, August 1992
 A>Description: Isolation and sequence of a sucrose synthase cDNA from developing rice
 A.Reference number: S25526
 A.Accession: S25526
 A.Molecule type: mRNA
 A.Residues: 1-190, 'P', 192-808 <ODE>
 A.Cross-references: EMBL:Z15028; NID:g20373; PIDN:CAA78747.1; PID:g20374
 C.Genetics:
 A.Introns: 32/2; 72/3; 123/2; 187/3; 227/2; 289/3; 331/3; 389/3; 428/3; 484/2; 559/2; 66
 C.Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
 C.Keywords: glycosyltransferase; hexosyltransferase
 F:276-750/Domain: sucrose-phosphate synthase homology <SSPS>

Query Match 76.7%; Score 3259.5; DB 2; Length 808;
 Best Local Similarity 76.2%; Pred. No. 4.3e-208;
 Matches 609; Conservative 91; Mismatches 96; Indels 3; Gaps 2;

Qy 6 LTRVHSLRERLDELTLAHRNEILALSRLEGKGLIHHQIILFEFAIPEENRKLKANG 65
 Db 5 LARHSLRERLDELTLAHRNEILALSRLEGKGLIHHQIILFEFAIPEENRKLKANG 62
 Qy 66 AFPEVLKASQEAIVLPVWALAVRRPGWMEIRVNVHVALVEELTVAAEYLFKELVDG 125
 Db 63 PFEDILRAQEAIVLPVWALAVRRPGWMEIRVNVHVALVEELTVAAEYLFKELVDG 122
 Qy 126 SSNGFVLELDEPPNSFPRTLSKISGNGVEFLNRHLSAKLPHDKSMPLLEFLRVH 185
 Db 123 HTNSFVLELDEPPNSFPRTLSKISGNGVEFLNRHLSAKLPHDKSMPLLEFLRVH 182
 Qy 186 CHKGNNMLNDRIQNLALQHVLRKAEEYLGTLPEPTCAEFENRFOEIGLGRGMDTAE 245
 Db 183 NHKGTMMNLNDRIQNLALQHVLRKAEEYLGTLPEPTCAEFENRFOEIGLGRGMDTAE 242

Qy 246 RVLEMIQLDLLLEATDPCLEKFLGRIPMFVNVILTPHGYFAODNVLYGYPDTGGQVY 305
 Db 243 RVLDITLHLLDLLEAPDPALEKFLGRTIPMFVNVILTPHGYFAODNVLYGYPDTGGQVY 302
 Qy 306 ILDOVRALENEMLRIRIKOQGLNTPRILITRLLPDAVGTTCQGLREKVGTEHSDILRV 365
 Db 303 ILDOVRALENEMLRIRIKOQGLNTPRILITRLLPDAVGTTCQGLREKVGTEHSDILRV 362
 Qy 366 PPRTEKIGVIRKVISFEKVPYLETETDVVAHEISKELHGPDLITGXSDGNIYASLLA 425
 Db 363 PPRTEKIGVIRKVISFEKVPYLETETDVVAHEISKELHGPDLITGXSDGNIYASLLA 421
 Qy 426 HKLGVTOCTIAALEKTYPPSDIYWKLEJRKVHPSQCFADLPAMNHTDITSTFOEI 485
 Db 422 HKLGVTOCTIAALEKTYPPSDIYWKLEJRKVHPSQCFADLPAMNHTDITSTFOEI 481
 Qy 486 AGSKDTQGYESHAFLLPGLYRVVHGIDVDPKFNIVSPGADMSIYEPYEEKRLKHF 545
 Db 482 AGSKDTQGYESHAFLLPGLYRVVHGIDVDPKFNIVSPGADMSIYEPYEEKRLKHF 541
 Qy 546 HPEIIDLTYTVENEHLCVANDRNKPLFTMPPLDRYKNTGLVWNGKPKARELANL 605
 Db 542 HPEIIDLTYTVENEHLCVANDRNKPLFTMPPLDRYKNTGLVWNGKPKARELANL 601
 Qy 606 VVVGDDRRKESKDLSEKAKMKMFELIDKYNLMGOFMWISSQMNRIKRVNLYRYICDTKG 665
 Db 602 VVVGDDRRKESKDLSEKAKMKMFELIDKYNLMGOFMWISSQMNRIKRVNLYRYICDTKG 661
 Qy 666 AFVQPALYEAAGLTIVBEAMTGLPTFATCNGGPAEIIYHGKSGFNIDPYHQDQADILVD 725
 Db 662 AFVQPALYEAAGLTIVBEAMTGLPTFATCNGGPAEIIYHGKSGFNIDPYHQDQADILVD 721
 Qy 726 FPEKCKDPSHMDKISQGLKRIEKKYTWKLYSERLLTLTGVGFWKVSUMLERRETRY 785
 Db 722 FPEKCKDPSHMDKISQGLKRIEKKYTWKLYSERLLTLTGVGFWKVSUMLERRETRY 781
 Qy 786 LEMFYALKYRKLAESVPLA 804
 Db 782 LEMFYALKYRKLAESVPLA 800

RESULT 8

T14338
 sucrose synthase (EC 2.4.1.13) isoform II - carrot
 CSpecies: Daucus carota (carrot)
 CDate: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 CAccession: T14338
 R.Stum, A.; Leinhardt, S.; Schatt, S.; Hardeggar, M.
 Plant Mol. Biol. 39, 349-360, 1999
 A>Title: Tissue-specific expression of two genes for sucrose synthase in carrot (Daucus
 A.Reference number: Z17990; MUID:99178785; PMID:10080700
 A.Accession: T14338
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-801 <STU>
 A.Cross-references: EMBL:Y16091; NID:g2760540; PIDN:CAA76057.1; PID:g2760541
 A.Experimental source: cultivar Namtalse; mature leaves
 C.Genetics:
 A.Gene: Susy*Dc2
 A.Introns: 29/2; 71/3; 122/2; 186/3; 298/3; 330/3; 388/3; 427/3; 483/2; 558/2; 746/2; 79
 C.Function:
 A.Description: catalyzes the reversible cleavage of sucrose into UDP-glucose and D-fructo-
 C.Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
 C.Keywords: glycosyltransferase; hexosyltransferase
 F:275-749/Domain: sucrose-phosphate synthase homology <SSPS>

Query Match 76.3%; Score 3238.5; DB 2; Length 801;
 Best Local Similarity 76.1%; Pred. No. 1e-206;
 Matches 604; Conservative 85; Mismatches 104; Indels 1; Gaps 1;

Qy 11 SLRERLDELTLAHRNEILALSRLEGKGLIHHQIILFEFAIPEENRKLKANGAFPEV 70
 Db 7 SLRERLDELTLAHRNEILALSRLEGKGLIHHQIILFEFAIPEENRKLKANGAFPEV 66

Qy	71	LKASQEAIVLPWVALAARPRGWEYIRVNVHALVVEELTVAEYLHFKEELVDGSSNGN	130
Db	67	LNSAQEAIVCSPMIALAIRLRGVWEYIRLVNHLQVVELTPVLYLKEELVNASNGN	126
Qy	131	FVLELDEFPNNSPPRPPLSLKSGINGVEFLNRHLNLSAKLPHDKESHMLLEFLRVHCHGK	190
Db	127	FVLELDFAPFPAISIPRPPLTKSINGVEFLNRHLNLSAKMFOKDSMHPLLDLRLHHNGR	186
Qy	191	NNMLNDRIONNALOHVLRKAEVYGLTPREPSCAEFHRFOEIGLERGMGDIAERVLEM	250
Db	187	TLMLNRRQVYVNGLODIIIRIAGEVLSKLPSTPTYSDFEHKFOEIGFERGMGDIAERVLEM	246
Qy	251	IQLLLELLEATDPCTLEKFLGRIPMVFNVVILTPHGFADQNVLAGPYDGGQVYIILQV	310
Db	247	FHMLLDLEAPDACLTEFLGKIPMI FNVVILSPHGFADQNVLAGPYDGGQVYIILQV	306
Qy	311	PALENNMLLRIRIKOQGLNTPRLLITRLIPDAVGTTCGQRLKRYGTHSDILKVPFTE	370
Db	307	PAMEREMTKRIKEQGLDIPRLLITRLIPDAVGTTCNLRLKRYGASHSHILRVPFTE	366
Qy	371	KGIVAKWISRPFKWPLYETVEDVAHESKELHGTPLIIGNXSDGNIVASLAKLGV	430
Db	367	KGILKMWISRE-VMPYMETFEDVAKELELKAKPDLIIGNVSEGNLVASLANKLGV	425
Qy	431	TOCTIAHALLEKTKYPSDSIYWKLEEDKXHFSCQFADLPFAMNHTDPIITSTFOELIAGSKD	490
Db	426	TOCTIAHALLEKTKYPSDSIYWKLEEDKXHFSCQFADLIAMNHTDPIITSTFOELIAGSKD	485
Qy	491	TVGQYESTATFLPGLYRVNHGIDVDPKFNIVSGADMEIYFPYTEERKRIKHPHIE	550
Db	486	TVGQYESTATFLPGLYRVNHGIDVDPKFNIVSGADTSVYYPYTEERKRRIALAHPIE	545
Qy	551	DLVLYKVENBEHLCVLNRNKPILFTMPRLDRVKULTGLVEMCGNPKRLRLANLVVVG	610
Db	546	DLPSVSVENKEHICVLKRYKPIFLTMARLDNVKULTGLVENYAKNPKRLRLANLVVVG	605
Qy	611	DRRKESKDLLEKAEKMMFELIDKYNLNGQFPMISSQNMNRIRANVELYRYICDTKGAFPQ	670
Db	606	DRRKESKDLLEKAEKMMFELIDKYNLNGQFPMISQAQNRVANGLERYICADTKGAFPQ	665
Qy	671	ALYEAFLGLTVVEAMTCGLPTFATCNGCPAEIIVHGKSGFNIDPHYGDQADILVDFEKC	730
Db	666	AFYEAFGLTVLIEAMTCGLPTFATIHGCPAEIIVHGTSGFHIDPHYGECAAEILVDFEKC	725
Qy	731	KKDPHMDKISQGLKRIEKKYTWKYSERLLTLGVYGFMKVGNLRRRSRRLYLEMY	790
Db	726	KTEPHMETISAGLKRIOEKITWQIYSERLTLGGVIGFMKXSKLDRIEIRYLEMFC	785
Qy	791	ALKYRKLAESVPLA 804	
Db	786	ALKYRNLAESVPLA 799	

RESULT 9
S32451
sucrose synthase (EC 2.4.1.13) Ss2 - barley

C:Species: Hordeum vulgare (barley)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Jun-1999
C:Accession: S32451; S31501
R:Martinez de Ilarduya, O.; Vicente-Carbajosa, J.; Sanchez de la Hoz, P.; Carbonero, P.
FEBS Lett. 320, 177-181, 1993
A:Title: Sucrose synthase genes in barley. cDNA cloning of the Ss2 type and tissue-specific
A:Reference number: S32451; MUID:93209382; PMID:9458435

A:Accession: S32451
A:Molecule type: mRNA

A:Residues: 1-816 <MAR>
A:Cross-references: EMBL:X69931; NID:g19099; PIDN:CAA49551.1; PID:g19100

C:Superfamily: sucrose synthase; sucrose-phosphate synthase homology
C:Keywords: glycoprotein; glycosyltransferase; hexosyltransferase

F:284-755/Domain: sucrose/sucrose-phosphate synthase homology <SSFS>
F:106,129,146,417,589,640/Binding site: carbonylrate (asn) (covalent) #status predicted

Query Match	75.9%;	Score 3222;	DB 2;	Length 816;
Best Local Similarity	75.0%;	Pred. No. 1.3e-205;	Mismatches 604;	Conservative 91;
			Mismatches 108;	Indels 2;
			Gaps 2;	
Qy	3	ERALTIVHSRERLDETLIARNEIILALLSIECKGIIQHHQIILEFE-AIPDENRK	61	
Db	7	ERALTIVHSRERIRGHSLAHTNELVAFSLRVNOGQMLDPOHITAEYNAIIPAEERK	66	
Qy	62	LANGAFEEVLKASQEAIVLPWVALAARPRGWEYIRVNVHALVVEELTVAEYLHFKEE	121	
Db	67	LKNTPFDLRLGQEAIVLPWVALAIRPRGWEYIRVVNSSELGVELSLRLYLQFREQ	126	
Qy	122	LVDSSNGNFVLELDEPEPNSSPPRPPLSLKSGINGVEFLNRHLNLSAKLPHDKESHMLLEF	181	
Db	127	LANGSTNNFVLELDFGPFNASPPRPPLSLKSGINGVOFLNRHLNLSKLFHDESNYPLNLF	186	
Qy	182	LRVCHGKNMMLNDRIONNALOHVLRKAEVYGLTPREPSCAEFHRFOEIGLERGMG	241	
Db	187	LRAHYNGKMTMMLNDRIRSLGTLQALRKAEHLISGLPADTPYEFHRRFOELGLEKMG	246	
Qy	242	DTAERVLEMIQLLLELLEATDPCTLEKFLGRIPMVFNVVILTPHGFADQNVLAGPYDGG	301	
Db	247	DCAQASETHILLDLLEAPDPSLEKFLGITIPVLANVILSPHGFADQNVLAGPYDGG	306	
Qy	302	QVYIILQVVALENEMLRIRIKOQGLNTPRLLITRLIPDAVGTTCGQRLKRYGTHSD	361	
Db	307	QVYIILQVVRAMENEMLRIRIKOGLDITPKLITRMLPDHNGTTCGQRLKRYGTHSD	366	
Qy	362	ILRVPRTEKGIYAKWISRPFKWPLYETVEDVAHESKELHGTPLIIGNXSDGNIVA	421	
Db	367	ILRVPFKEIDGIVAKWISRE-VMPYLAETVDVAHELISGELQANPDLIIGNVSDGNIVA	425	
Qy	422	SILAHKLVNTOCTIAHALLEKTKYPSDSIYWKLEEDKXHFSCQFADLPFAMNHTDPIITST	481	
Db	426	CLIAHKLVNTHCTIAHALLEKTKYPSDLYWKXFDHNFSCQFADLIAMNHTDPIITST	485	
Qy	482	FOETIAGSKDTVGQYESTATFLPGLYRVNHGIDVDPKFNIVSGADMEIYFPYTEERK	541	
Db	486	FOETIAGNKDTVGQYESHMAFTMPGLYRVNHGIDVDPKFNIVSGADMSIFPYTEQOKR	545	
Qy	542	LKHHPREIIDLTVKVENBEHLICVLNRNKPILFTMPRLDRVKULTGLVEMCGNPKRLRE	601	
Db	546	LTSIHTEIIEELFSDVNEAEKFFVKDKKPIITSMARLDRVKULTGLVENYGRNPKRLRE	605	
Qy	602	LANLVVVGDRRKESKDLLEKAEKMMKFELIDKYNLNGQFPMISSQNMNRIRANVELYRYIC	661	
Db	606	LNVLVVVGDRGHGVSXDKQEBQVEFKMPDLIEKTNLSGHIMISQNMNRVANGLERYRYIC	665	
Qy	662	DTKGAFVQALYEAFLGLTVVEAMTCGLPTFATCNGCPAEIIVHGKSGFNIDPHYGDQAD	721	
Db	666	DMKGAFVQAPYEAFLGLTVLIEAMTCGLPTFATVAGCPAEIIVNGVSGYHIDPHYQDKASA	725	
Qy	722	ILVDFEKKCKDPHMDKISQGLKRIEKKYTWKYSERLLTLGVYGFWMKVSNLRRSR	781	
Db	726	LIVGFGKQCEBPSHMNKISQGLQRIEKKYTWKYSERLLTLGGVYGFWMKVSNLDRRE	785	
Qy	782	SRRLYLEMYALKYRKLAESVPLAE	806	
Db	786	TRRLYLEMYALKYRKMAATVPLAVE	810	

RESULT 10

S29242
sucrose synthase (EC 2.4.1.13) Ss1 - barley

N:Alternate names: sucrose-UDP glucosyltransferase; UDPglucose-fructose glucosyltransferase
C:Species: Hordeum vulgare (barley)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Jun-1999

C:Accession: S29242; S21494
R:de la Hoz, P.S.; Vicente-Carbajosa, J.; Mena, M.; Carbonero, P.

FEBS Lett. 310, 46-50, 1992
A:Title: Homologous sucrose synthase genes in barley (Hordeum vulgare) are located in ch

A:Reference number: S29242; MUID:92405741; PMID:1138123

A:Accession: S29242

A:Molecule type: mRNA
 A:Residues: 1-807 <HOZ>
 A:Cross-references: EMBL:X65871; NID:g19105; PIDN:CAA46701.1; PID:g19106
 C:Genetics:
 A:Map position: 7H
 C:Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:276-749/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 75.6%; Score 3209; DB 2; Length 807;
 Best Local Similarity 75.8%; Pred. No. 9,5e-205;
 Matches 606; Conservative 89; Mismatches 100; Indels 4; Gaps 3;

```

QY 6 LTRVSLRERLDETLAHNEITALLSRLEGKGILOHNOIILEFEALPEENRK- 65
DB 5 LTRLSLBERLGATSSHNELIALPSRVHOGKMLQHQLAEFDALFESDKERYA-- 62
QY 66 AFEEVLKASQEAIVLPWVALAVRPRGWEYIRVNVHVALVEELTVAEYLHFKELVDG 125
DB 63 PFEDILRAAGEAIVLPWVALAIRPTGVWDYIRVNVSLAVEELTVAEYLHFKELVDG 122
QY 126 SSNGNFVLELDFEPNNSPPRPTLSKISGVNVEFLNRHLSAKLPHDKSNMPLLEFLRVH 185
DB 123 HASRFVLELDFEPNNSPPRPTLSKISGVNVEFLNRHLSAKLPHDKSNMPLLEFLRVH 182
QY 186 CHKGKMMMLNDRIQNLNLQHLARKAEYLGTLPPETPCAEFEHRFOEIGLERGMGDTAE 245
DB 183 NKGKMTMLNDRIQNLNLQHLARKAEYLGTLPPETPCAEFEHRFOEIGLERGMGDTAE 242
QY 246 RYLEMIQILLDLLEATPCTLEKFLGRIPMVFNVLITPHGYFAODNVLGYPDTGGGVY 305
DB 243 RVHDTIHLDDLEAPDPSLEKFLGRIPMVFNVLITPHGYFAODNVLGYPDTGGGVY 302
QY 306 ILDOVRALBEMMLIRIKOQGLNITPRILITRLLPDVAGTTGGRLEKYGTEHSDILRV 365
DB 303 ILDOVRALBEMMLIRIKOQGLNITPRILITRLLPDVAGTTGGRLEKYGTEHSDILRV 362
QY 366 PPRTEKGIYVKMISRFKWPYLETYEDVAHEISKELHGTPLDIIGNXSDGNIVASILA 425
DB 363 PPRTEKGIYVKMISRFKWPYLETYEDVAHEISKELHGTPLDIIGNXSDGNIVASILA 420
QY 426 HKLGVTGCTIAHALEKTKYPSDSIYMKLEDKYHFSQFTADLFAMNHTDPIITSTPOEI 485
DB 421 HKLGVTGCTIAHALEKTKYPSDSIYMKLEDKYHFSQFTADLFAMNHTDPIITSTPOEI 480
QY 486 AGSKDTVGQYESHAFPLPGLYRVVHGIDVDFPKFNIVSPGADMEIYPPYTEEKRLKH 545
DB 481 AGSKDTVGQYESHAFPLPGLYRVVHGIDVDFPKFNIVSPGADMEIYPPYTEEKRLKH 540
QY 546 HPEIEDLLYTKVNEHEHLCVLDNRKPIITMPRLDRVKNLTGLVEMGCKNPKLELANL 605
DB 541 HSEIELLYSDVENDEHKFVLDKDRNKPITFSMARLDVKNMTGLVEMGCKNPKLELANL 600
QY 606 VVVGDRRKESKDLBEKAKEMKMFELIDKYNLNGOFWRISQMRIRVVELYRYICDTKG 665
DB 601 VIVAGDHGKESDRBEQAEFKRMYSILBEYKLGKIHRIISAGMNVKRGELRYRICDTKG 660
QY 666 AFVOPALYEAGLTVVEAMTCGLPTFATCNGGPAEIIIVHSGSGFNIDPYHGDQADII 725
DB 661 AFVOPALYEAGLTVVEAMTCGLPTFATCNGGPAEIIIVHSGSGFNIDPYHGDQADII 720
QY 726 FFEKCKDPSSHMDKISQGLKRIEKTWKIYSEBLTLTGYYGKWKVSNLERESRY 785
DB 721 FFEKCKDPSSHMDKISQGLKRIEKTWKIYSEBLTLTGYYGKWKVSNLERESRY 780
QY 786 LEMFALAKYRKLAESVPLA 804
DB 781 LEMFALAKYRKLAESVPLA 799

```

RESULT 11
 T06497
 probable sucrose synthase (EC 2.4.1.13) 2 - garden pea

C:Species: Pisum sativum (garden pea)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
 C:Accession: T06497
 R:Buchner, P.
 submitted to the EMBL Data Library, October 1997
 A:Reference number: Z15720
 A:Accession: T06497
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-809 <BUC>
 A:Cross-references: EMBL:AJ001071; PIDN:CAA04512.1
 A:Experimental source: cultivar Filson
 C:Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:282-757/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 70.3%; Score 2985; DB 2; Length 809;
 Best Local Similarity 70.4%; Pred. No. 6,9e-190;
 Matches 565; Conservative 112; Mismatches 120; Indels 6; Gaps 4;

```

QY 7 TRVSLRERLDETLAHNEITALLSRLEGKGILOHNOIILEFEALPEENRK--KLAN 64
DB 8 TRVPSIRDVODTLAHNEELISLSRYVAGKGILOPHNIDDELNLIGSDHATLDLKN 67
QY 65 GAFFEVKASQEAIVLPWVALAVRPRGWEYIRVNVHVALVEELTVAEYLHFKELVDG 124
DB 68 GFPGIINSQEAIVLPWVALAVRPRGWEYIRVNVHVALVEELTVAEYLHFKELVDG 127
QY 125 SSNGNFVLELDFEPNNSPPRPTLSKISGVNVEFLNRHLSAKLPHDKSNMPLLEFLRV 184
DB 128 GSKNDNITLEDLDFEPNNSPPRPTLSKISGVNVEFLNRHLSAKLPHDKSNMPLLEFLRV 187
QY 185 HCHKGKMMMLNDRIQNLNLQHLARKAEYLGTLPPETPCAEFEHRFOEIGLERGMGDTA 244
DB 188 HCHKGKMMMLNDRIQNLNLQHLARKAEYLGTLPPETPCAEFEHRFOEIGLERGMGDTA 247
QY 245 ERLVEMIQILLDLLEATPCTLEKFLGRIPMVFNVLITPHGYFAODNVLGYPDTGGGVY 304
DB 248 ARVLEMMMLDDIIOAPDPSLEKFLGRIPMVFNVLITPHGYFAODNVLGYPDTGGGVY 307
QY 305 YILDQVRALBEMMLIRIKOQGLNITPRILITRLLPDVAGTTGGRLEKYGTEHSDILRV 364
DB 308 YILDQVRALBEMMLIRIKOQGLNITPRILITRLLPDVAGTTGGRLEKYGTEHSDILRV 367
QY 365 VPRTEKGIYVKMISRFKWPYLETYEDVAHEISKELHGTPLDIIGNXSDGNIVASILA 424
DB 368 VPRTEKGIYVKMISRFKWPYLETYEDVAHEISKELHGTPLDIIGNXSDGNIVASILA 426
QY 425 AHKLGVTGCTIAHALEKTKYPSDSIYMKLEDKYHFSQFTADLFAMNHTDPIITSTPOE 484
DB 427 AHKLGVTGCTIAHALEKTKYPSDSIYMKLEDKYHFSQFTADLFAMNHTDPIITSTPOE 486
QY 485 IAGSKDTVGQYESHAFPLPGLYRVVHGIDVDFPKFNIVSPGADMEIYPPYTEEKRLKH 544
DB 487 IAGSKDTVGQYESHAFPLPGLYRVVHGIDVDFPKFNIVSPGADMEIYPPYTEEKRLKH 546
QY 545 FHEIEDLLYTKVNEHEHLCVLDNRKPIITMPRLDRVKNLTGLVEMGCKNPKLELANL 604
DB 547 FHEIEDLLYTKVNEHEHLCVLDNRKPIITMPRLDRVKNLTGLVEMGCKNPKLELANL 606
QY 605 LVVVG--DRRKESKDLBEKAKEMKMFELIDKYNLNGOFWRISQMRIRVVELYRYICD 662
DB 607 LVVVG--DRRKESKDLBEKAKEMKMFELIDKYNLNGOFWRISQMRIRVVELYRYICD 665
QY 663 TKGAFFVOPALYEAGLTVVEAMTCGLPTFATCNGGPAEIIIVHSGSGFNIDPYHGDQADI 722
DB 666 TKGAFFVOPALYEAGLTVVEAMTCGLPTFATCNGGPAEIIIVHSGSGFNIDPYHGDQADI 725
QY 723 LVDFPEKCKDPSSHMDKISQGLKRIEKTWKIYSEBLTLTGYYGKWKVSNLERES 782
DB 726 LVDFPEKCKDPSSHMDKISQGLKRIEKTWKIYSEBLTLTGYYGKWKVSNLERES 785
QY 783 RRYLEMFALAKYRKLAESVPLAE 805

```

Db 766 RRYLEMFYILKFRDIANSVPFLAK 808

||||| | : | ||||| :
RESULT 12
B85029
Probable sucrose synthetase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001
C:Accession: B85029
Randommous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Harbor Laboratory Press
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617158
A:Accession: B85029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-809 <STO>
A:Cross-References: GB:NC_001268; NID:97268988; PIDN:CAB80721.1; GSFDB:GN00140
C:Gene: ATG02280
A:Map position: 4
C:Superfamily: sucrose synthase; sucrose/sucrose-phosphate synthase homology

Query Match 69.7%; Score 2962; DB 2; Length 809;
Best Local Similarity 69.3%; Pred. No. 2,3e-188;
Matches 558; Conservative 112; Mismatches 131; Indels 4; Gaps 3;

1 MAERLTVHSLRELDLTLLAHREILLALSRIGKGKGLIHHQILLEPAIL--PEN 58
1 MANPKLTRVLSLRDVRQDTLSAHNELVALSRYYDGKGIQPNNLIDELSEVIGDET 60
59 RKKLNAFFFEVLTKASOEALVLPWVALAVRPPGWMEYIRNVHALVVEELTYAEYLHF 118
61 KKSLSDEGGELTSGAMEALIVPPFVALAVRRPEWMEYVANNVELSVDQLTSEYLRF 120
119 KEELYDSSNGNFVLELDFEPNSSFPRPTLSKISGNGEFLNRHLAKLTHDKSNMPL 178
121 KEELVDGPNSDFCLELDFEPFNANVPRPSRSSISGNGVOFLNRHLSVMFRNKDCLEPL 180
179 LEFLAVCHKGKMMLNDRIQNLMNLQHVLRPAEEYLGTLPPETPCAFEFHRFOEIGER 238
181 LDFLVHKYKGMPLMNDIQISIRLQQLSAEDHISIKLSQEITPFSEFYALQGMPFK 240
239 GWGDTFAERYLEMIQLDLLEATDPTCLEFKRIPMWPNVILTPHSYFADQNVLYGYPD 298
241 GMGDTAGRVLEMNHLSLDLQAPDSLSLEKFGWPMFVNVIILSPHYPOGANVLYGPD 300
299 TGGGVVVYIIDQVRALENEMLRKKOQGLNTPRILITRLRPDAVGTTGGORLEKYGTGE 358
301 TGGVVVVYIIDQVRALETEMLRIRIKQGLDISPSIIIVTRLIPDAKITCNORLERVSGETE 360
359 HSDILARPRETEKGVRRKISRFEKWVPYLETYTEDVAHEISKELHGPRDILINGXSDGN 418
361 HHHLIRVPRSEKGLRKWISRFD--VWPYLENVADDAASEIVGELQGVDPFLIGNYSNGN 419
419 IYASLIHLNLTGTOCTIAHALETKEYPDSDIYWKKLDEKYHSCOPTADLEPANHTDPII 478
420 LVASLMARMGSTOCTIALALEKTYPDSDIYWKOPDNKYHSCOPTADLAMNADPII 479
479 TSTFOEIASKDVTQGYESHTAFTLPGLYRVVHGIDVDPKFNISPADMEIFYPYTEE 538
480 TSTYOELINQTNVQGYESHGAFTLPGLYRVVHGIDVDPKFNISPADMTIYPPYSEE 539
539 KARLKHPHEIEDLYTKVENEHELCVLNDRKPLFLTFMRPLDRKYNLTGLVENCGKNPK 598
540 TRRLTAHGISIEBMLYSPQTDEHVTLSDRSKPLIFLSMARLDKYNKISGLVEMYSKNTK 599
599 LRRLNLTVVVGD--RRKESKDLEEKAEMKMKELLIDKNVNLNQPMNISOMNRIRINVELY 657
600 LRRLNLTVAAGNIDVYKSKDRBEIVEIKKMNKLNKYGLDOGFMMITATQNRARNGLY 659
658 RRICOTKGAFOVPALYEAFGLTVTEAMTCGLPTFACNGCPAAIIVHKSGENIDPYHD 717

[illegible]

Db 147 SSSINGVQVUNRHUSIMFNKESMEPLLEPLRTHKIDGRPMINDRIQNIPIIQGALA 206
Qy 210 KAEELYGLTPPETPCAEFEHREFOEIGLERGWDTAERYLEMIOQLLDLEATDPCTLEKF 269
Db 207 RAEEELSLPLATPYSEFEPELQGMFERGWDTAQKVSBNVHLDDILQAPDPSVLETF 266
Qy 270 LGRIMFNVNVLTPHGFPAQDNVLGYPDGQVYIILDQVPALENEMLRIKQOGLNIT 329
Db 267 LGRIMFNVNVLISRYGAQANVGLPDTGAQVYIILDQVPALENEMLRIKQOGLNIT 326
Qy 330 PRILITLPPAVGTGGORLEKXYGTEHSDILRVPRTEKGIKRWISREKWPYLE 389
Db 327 PRILITLPPAKGTCQRLERVSGETHAHLRIPRTEKGIKRWISREKWPYLE 385
Qy 390 TYTEDVAHEISKELHGTPLDILGNXSDGNIVASLNAHLGVTOCTIAHALEKTKYPSDI 449
Db 386 TPAEDASNEISAELOGVNLIIGNVSDGNVSLASKLGVIOCNIAHALEKTKYPSDI 445
Qy 450 YWKLEEDKYHSCQFTADLFAMNHTDFTITSTFOEIASGKDTVGQYESHTAFTLPGLYRV 509
Db 446 YWRNHEDKXHFSSQFTADLIAMNNADFITSTYQEIAGSKNNVGQYESHTAFTMPGLYRV 505
Qy 510 VHGIVDPKFNIVSPGADMEIYFPYTEKRLKHFPHEIDLVTKVNEHEHLCVINDR 569
Db 506 VHGIVDPKFNIVSPGADMTIYFPYSDKERRLTALHSEIEHLFSAQONDEHVGLSIQ 565
Qy 570 NKPIFTMPRLDRVKNLTGLVEMCGKNPKYRELANLVVVG-DRRKESKDLSEKAEKMM 628
Db 566 SKPIIFSMARLDRVKNLTGLVEMCGKNPKYRELANLVVVGIDENQSRDEMAEIOKM 625
Qy 629 FELIKYNLNGOFNRISSOMNRIINVELYRYICDTKGAFOPALYEAFTVVEAMTCGL 688
Db 626 HSLIEQYDLHGEFRWIAAQMNVRNGELYRYIADTKGVFOPAFYEAFTVVEAMTCAL 685
Qy 689 PTFATONGGPAEIIYHSGSEFNIDPYHGOADILVDFEKKCKDPSHMDKISOGGLKRI 748
Db 686 PTFATCHGGPAEIIENGVSGFHIIDPYHDPQVAGSLA-LFETCNINPNHVKISBSGLKRI 744
Qy 749 BEKYTWKIYSERLTLTGYYGFMKIVSNLERRESRYLEMFYALKYRKLAESVPLAE 806
Db 745 YERYTWKYSERLTLGAGYAFMKIVSKLERRETRYLEMFYSLKFRDLANSIPLATD 802

Search completed: June 2, 2004, 14:53:14
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model1

Run on: June 2, 2004, 14:47:54 ; Search time 18 Seconds

(without alignments)
2331.586 Million cell updates/sec

Title: US-10-003-405-2

Sequence: 1 MAERALTRVHSLRERLDETL.....EMFYALKRKLAESVPLAE 806

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3763.5	88.6	805	1	SUSY SOYBN
2	3714.5	87.5	805	1	SUSY PHAVN
3	3669.5	86.4	805	1	SUSY VICFA
4	3648.5	85.9	806	1	SUSY MEDSA
5	3589.5	84.5	806	1	SUSY ARATH
6	3515.5	82.8	803	1	SUSY ALNGL
7	3501.5	82.4	805	1	SUSY SOLTU
8	3496.5	82.3	805	1	SUSY LYCES
9	3476.5	81.9	805	1	SUS2 SOLTU
10	3400.5	80.1	808	1	SUS1 DAUCA
11	3301	77.7	820	1	SUS2 TULGE
12	3295	77.6	816	1	SUS2 MAIZE
13	3273	77.1	805	1	SUS1 TULGE
14	3273	77.1	816	1	SUS2 ORYSA
15	3268.5	77.0	802	1	SUS1 MAIZE
16	3262	76.8	816	1	SUS3 ORYSA
17	3252.5	76.6	808	1	SUS1 ORYSA
18	3238.5	76.3	801	1	SUS2 DAUCA
19	3222	75.9	816	1	SUS2 HORVU
20	3209	75.6	807	1	SUS1 HORVU
21	2985	70.3	809	1	SUS2 PEA
22	2865.5	67.5	805	1	SUS2 ARATH
23	2841	66.9	766	1	SUSY BETTU
24	661	15.6	218	1	SUSY SACOF
25	437	10.3	1059	1	SUS VICFA
26	436	10.3	1068	1	SUS MAIZE
27	431	10.1	1056	1	SUS SPIOL
28	421.5	9.9	1045	1	SUS2 BETTU
29	418.5	9.9	1081	1	SUS2 CRAPL
30	415.5	9.7	1053	1	SUS SOLTU
31	407	9.6	1057	1	SUS CITUN
32	394	9.3	1054	1	SUS1 CRAPL
33	387.5	9.1	1049	1	SUS ORYSA

34	156	3.7	377	1	YPJH_BACSU
35	148	3.5	5430	1	MACF_HUMAN
36	147	3.5	5938	1	MACF_HUMAN
37	142	3.3	358	1	YC07_KLEPN
38	135.5	3.2	406	1	YAS9_METUA
39	131	3.1	5327	1	MACF_MOUSE
40	129	3.0	390	1	YG07_METUA
41	128	3.0	5171	1	BPEA_HUMAN
42	127	3.0	380	1	CAPM_STRAU
43	124	2.9	359	1	YOGM_BACSU
44	124	2.9	461	1	GP13_YEAST
45	123	2.9	1042	1	TIH1_METUA

ALIGNMENTS

RESULT 1
ID SUSY SOYBN STANDARD; PRT: 805 AA.
AC P13708; Q22624;
DT 01-JAN-1990 (Rel. 13, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase)
DE (Nodulin-100).
GN GN
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Euphorbia I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Hobbit; TISSUE=Root nodules;
RA Zhang X.-Q., Verma D.P.S., Patil S., Arredondo-Peter R., Miao G.-H.,
RA Kuismanen R., Klucac R.V., Chollet R.;
RT "Cloning of a full-length sucrose synthase cDNA from soybean (Glycine
RT max) root nodules."
RL (In) Plant Gene Register PGR97-173.
RN [2]
RP SEQUENCE OF 679-805 FROM N.A.
RC STRAIN=cv. Prize; TISSUE=Root nodules;
RX MEDLINE=88033030; PubMed=2889731;
RA Thummel F., Verma D.P.S.;
RT "Nodulin-100 of soybean is the subunit of sucrose synthase regulated
RT by the availability of free heme in nodules."
RL J. Biol. Chem. 262:14730-14736 (1987).
RN [3]
RP PHOSPHORYLATION.
RX MEDLINE=97379292; PubMed=9237614;
RA Zhang X.-Q., Chollet R.;
RT "Seryl-phosphorylation of soybean nodule sucrose synthase
RT (nodulin-100) by a Ca²⁺-dependent protein kinase."
RL FEBS Lett. 410:126-130 (1997).
CC -!- FUNCTION: Sucrose-leaving enzyme that provides UDP-glucose and
CC fructose for various metabolic pathways.
CC -!- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
CC -!- SUBUNIT: Homotrimer.
CC -!- PTM: Phosphorylated on serine residue(s).
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
CC sucrose synthase subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC EMBL: AF030231; AAC39323.1; -.
CC PIR: A29484; A29484.

DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR000368; Sucrose synth.
 DR Pfam; PF00534; Glycosyltransferase_1; 1.
 DR Pfam; PF00862; Sucrose synth. 1.
 KM Transferrase; Glycosyltransferase; Nodulation; Phosphorylation.
 FT MOD_RES 11 11 PHOSPHORYLATION (B/CDPK) (POTENTIAL).
 FT CONFLICT 680 680 V -> L (IN REF. 2).
 FT CONFLICT 737 738 DK -> ET (IN REF. 2).
 FT CONFLICT 804 804 A -> V (IN REF. 2).
 SQ SEQUENCE 805 AA; 92243 MW; BBI4E5E8B4F5587B CRC64;

Query Match 88.6%; Score 3763.5; DB 1; Length 805;
 Best Local Similarity 88.0%; Pred. No. 8.6e-235;
 Matches 709; Conservative 49; Mismatches 47; Indels 1; Gaps 1;

QY 1 MAERALTIVHSRLRDELTLAHNEILLALSRIEGKGGILQHHQIILFEALPEENRK 60
 DB 1 MATRLTRVHSRLRDELTLAHNEILLALSRIEGKGGILQHHQIILFEALPEENRK 60
 QY 61 KLANGAFEEVLKASQEAIVLPWVALAVPRPGVMEYIRVNVHALVVEELTVAEYLHKE 120
 DB 61 KLTDAFGEVLRSTGEAIVLPWVALAVPRPGVMEYIRVNVHALVVEELQPAEYLHKE 120
 QY 121 ELVDGSSNGNFVLELDEPPFNSSPPRPTLSKISNGVEFLNRHLSAKLPHDKESMHPLE 180
 DB 121 ELVDGSSNGNFVLELDEPPFNAPRPTLNKISNGVQFNLNRHLSAKLPHDKESMHPLE 180
 QY 181 FLRVCHGKXMMNLNDRIQNLALOHVLRKAEYVGLTPREPCAEFEHRFOELGLEBGM 240
 DB 181 FLRLSVGKXKMLNDRIQNPALQHVLRKAEYVGLTPRETPSEFEHRFOELGLEBGM 240
 QY 241 GDTAEVLEMIQLLDLLEATDPTCLEKFLGRIPMVFNVLITPPGYFAQDNVLGYPDGT 300
 DB 241 GDNARVLESQLLDLLEAPDPTCLEFLGRIPMVFNVLITSPGYFAQDNVLGYPDGT 300
 QY 301 GGVVYIILDOVRALENEMLRIRKQOGLNITPRILITRLPDAVGTTCGORLEKYGTBHS 360
 DB 301 GGVVYIILQVRALENEMLRIRKQOGLNITPRILITRLPDAVGTTCGORLEKYGTBHS 360
 QY 361 DILRVPRTEKGIYKWKISRPEKWPYLETYTEDEVVAHSEIKELHOTPDLIIGNSDGNIV 420
 DB 361 HILRVPRTEKGIYKWKISRPE-VMPLYETYTEDEVVAHSEIKELQKPDILIGNSDGNIV 419
 QY 421 ASLNAHLKLVTOCTIAHALEKTKYPSDSIYWKLEDKYHFSQCFADLPFAMNHTDFIITS 480
 DB 420 ASLNAHLKLVTOCTIAHALEKTKYPSDSIYWKLEBRVHFSQCFADLPFAMNHTDFIITS 479
 QY 481 TFOEJAGSKDTVGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSGADMEIYEPYTEER 540
 DB 480 TFOEJAGSKDTVGOYESHTAFTLPGLYRVVHGIDVDPKFNIVSGADQITVFPPTETSR 539
 QY 541 RLKHHPREIEDLLYTKVNEEHLCVLNPBNKPILTMPRLDVKVLGTGVEMCGNPKLR 600
 DB 540 RLTSHPREIEDLLYTSVNEEBHICVLKDSKPIITTMALDVKVITGTVVEYGNNAKLR 599
 QY 601 ELIANLVVGGDRRKSSKOLEEKAEMKMEFELIDKTNLNGOFRRMISSQMRIRNVELYRYI 660
 DB 600 ELVNVLVVAGDRRKSSKOLEEKAEMKMYGLLETNKLNGOFRRMISSQMRIRNVELYRYI 659
 QY 661 CDTKCAFVQPALYEAFGLTVVEAMTCGLPTFATCNGPAEITVHKSQFNIDPYHGDAA 720
 DB 660 CDTKCAFVQPAVEAFGLTVVEAMTCGLPTFATCNGPAEITVHKSQFNIDPYHGDAA 719
 QY 721 DILVVFPECKCKDPBHMWKISOGGLKRIEKKYTWKYSERLLTLTGUVGPMKHSVNLRR 780
 DB 720 DILVVFPECKCKDPBHMWKISOGGLKRIEKKYTWKYSERLLTLTGUVGPMKHSVNLRR 779
 QY 781 ESRRYLEMFYALKYRKLAEVPLAE 806
 DB 780 ESRRYLEMFYALKYRKLAEVPLAE 805

RESULT 2

SUSY PHAAU
 ID SUSY PHAAU STANDARD; PRT; 805 AA.
 AC 001390;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase).
 GN S51.
 OS Phaseolus aureus (Mung bean) (Vigna radiata).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.
 OC NCBI_TaxID=3916;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 517-537.
 RA Arai M., Mori H., Imaeki H.;
 RT "Expression of the gene for sucrose synthase during growth of mung
 bean seedlings.";
 RL Plant Cell Physiol. 33:503-506 (1992).
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
 sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by, and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D10266; BAA01108.1; -;
 DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR000368; Sucrose synth.
 DR Pfam; PF00534; Glycosyltransferase_1; 1.
 DR Pfam; PF00862; Sucrose synth. 1.
 KM Transferrase; Glycosyltransferase.
 SQ SEQUENCE 805 AA; 92092 MW; A39683CFD54EFA4 CRC64;

Query Match 87.5%; Score 3714.5; DB 1; Length 805;
 Best Local Similarity 87.3%; Pred. No. 1.2e-221;
 Matches 704; Conservative 47; Mismatches 54; Indels 1; Gaps 1;

QY 1 MAERALTIVHSRLRDELTLAHNEILLALSRIEGKGGILQHHQIILFEALPEENRK 60
 DB 1 MATRLTRVHSRLRDELTLAHNEILLALSRIEGKGGILQHHQIILFEALPEENRK 60
 QY 61 KLANGAFEEVLKASQEAIVLPWVALAVPRPGVMEYIRVNVHALVVEELTVAEYLHKE 120
 DB 61 KLTDAFGEVLRSTGEAIVLPWVALAVPRPGVMEYIRVNVHALVVEELQPAEYLHKE 120
 QY 121 ELVDGSSNGNFVLELDEPPFNSSPPRPTLSKISNGVEFLNRHLSAKLPHDKESMHPLE 180
 DB 121 ELVDGSSNGNFVLELDEPPFNAPRPTLNKISNGVQFNLNRHLSAKLPHDKESMHPLE 180
 QY 181 FLRVCHGKXMMNLNDRIQNLALOHVLRKAEYVGLTPREPCAEFEHRFOELGLEBGM 240
 DB 181 FLRLSVGKXKMLNDRIQNPALQHVLRKAEYVGLTPRETPSEFEHRFOELGLEBGM 240
 QY 241 GDTAEVLEMIQLLDLLEATDPTCLEKFLGRIPMVFNVLITPPGYFAQDNVLGYPDGT 300
 DB 241 GDNARVLESQLLDLLEAPDPTCLEFLGRIPMVFNVLITSPGYFAQDNVLGYPDGT 300
 QY 301 GGVVYIILDOVRALENEMLRIRKQOGLNITPRILITRLPDAVGTTCGORLEKYGTBHS 360
 DB 301 GGVVYIILQVRALENEMLRIRKQOGLNITPRILITRLPDAVGTTCGORLEKYGTBHS 360
 QY 361 DILRVPRTEKGIYKWKISRPEKWPYLETYTEDEVVAHSEIKELHOTPDLIIGNSDGNIV 420
 DB 361 HILRVPRTEKGIYKWKISRPE-VMPLYETYTEDEVVAHSEIKELQKPDILIGNSDGNIV 419

QY 421 ASLNAHKLGVCTCTTAHALEKTKYPSDSIYWKKEBDKHFSCQFTADLFPANNHTDEIITS 480
 DB 420 ASLNAHKLGVCTCTTAHALEKTKYPSDSIYWKKEBDKHFSCQFTADLFPANNHTDEIITS 479
 QY 481 TFOEIASGSDTVGOYESHTAFTLPGLYRVVHGIDVDFPKFNIYVSGADMEIYFPYTEER 540
 DB 480 TFOEIASGSDTVGOYESHTAFTLPGLYRVVHGIDVDFPKFNIYVSGADMEIYFPYTEER 539
 QY 541 RLKHPHEIEDLLYTKVNEEHLVYNDNRKPIILFMPRLDVKVNLTLGVYCGNPKYR 600
 DB 540 RLKHPHEIEDLLYTKVNEEHLVYNDNRKPIILFMPRLDVKVNLTLGVYCGNPKYR 599
 QY 601 ELANLVVGGRRRKSKELEKAEKMKFELIDKYNLNGQFWMISSQNNRINVELYRYI 660
 DB 600 ELVNLVVAAGRRRKSKELEKAEKMKFELIDKYNLNGQFWMISSQNNRINVELYRYI 659
 QY 661 CDTKCAFQOPALYEAFTLVVAMTCGLPTPATCNGPAEIIYHGKSGFNIDPYHGDAA 720
 DB 660 ADTKCAFQOPAYEAFGLTVVAMTCGLPTPATCNGPAEIIYHGKSGFNIDPYHGDAA 719
 QY 721 DILVDFEKKCKDPBHMWKISQGLKRIEKKYTWKYSERLLTLTGYYGFMKGVNLEER 780
 DB 720 DILVDFEKKCKDPBHMWKISQGLKRIEKKYTWKYSERLLTLTGYYGFMKGVNLEER 779
 QY 781 ESRRLYEMFYALKYRKLAESVPLAE 806
 DB 780 ESRRLYEMFYALKYRKLAESVPLAE 805

RESULT 3
 SUSY VICFA STANDARD; PRT; 806 AA.

ID AC P31926: 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase).
 OS Vicia faba (Broad bean).
 GN Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Vicia.
 OX NCBI Taxid=3906;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Fribo; PubMed=7764025;
 RX MEDLINE=93379571; PubMed=7764025;
 RA Heim U., Weber H., Baumelein H., Mobus U.;
 RT "A sucrose-synthase gene of Vicia faba L.: expression pattern in
 RT developing seeds in relation to starch synthesis and metabolic
 RT regulation.";
 RL Planta 191:394-401 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Kline Thueringer; TISSUE=Root nodules;
 RX MEDLINE=94003420; PubMed=8400379;
 RA Kuether H., Fruhling M., Perlick A.M., Puchler A.;
 RT "The sucrose synthase gene is predominantly expressed in the root
 RT nodule tissue of Vicia faba.";
 RT Mol. Plant Microbe Interact. 6:507-514 (1993).
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 CC fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
 CC sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC -----
 DR EMBL: X69773; CAA9428.1; -.
 DR EMBL: M97551; AAC37346.1; -.
 DR PIR: S31479; S31479.
 DR InterPro: IPR001296; Glyco_trans_1.
 DR InterPro: IPR000368; Sucrose_synth.
 DR Pfam: PF00534; Glycos_transf_1; 1.
 DR Pfam: PF00862; Sucrose synth; 1.
 KW Transferase; Glycosyltransferase.
 SQ SEQUENCE 806 AA; 92520 MM; A51B004C0732F306 CRC64;

Query Match 86.4%; Score 3669.5; DB 1; Length 806;
 Best Local Similarity 86.2%; Pred. No. 9.8e-229;
 Matches 695; Conservative 52; Mismatches 58; Indels 1; Gaps 1;

QY 1 MAERALTVHSGIRELDETLAHRNEILLASRIEKGKGILOHQTILEEPAIPEERK 60
 DB 1 MATERTLVHSGIRELDETLAHRNEILLASRIEKGKGILOHQTILEEPAIPEERK 60
 QY 61 KLANGAFEVYKASGEATVLPWVALAVRPPGWWEYIRVNVHALLVVELTVAEYLHFK 120
 DB 61 KLTDGAFGEVLRSTGEATVLPWVALAVRPPGWWEYIRVNVHALLVVELTVAEYLHFK 120
 QY 121 ELVDSSNGNFVLELDFEPFNSSPFRPTLSKISGVEFLNRHLSAKLFHDKESMHPLE 180
 DB 121 ELVDSSNGNFVLELDFEPFNSSPFRPTLSKISGVEFLNRHLSAKLFHDKESMHPLE 180
 QY 121 ELVDSSNGNFVLELDFEPFNSSPFRPTLSKISGVEFLNRHLSAKLFHDKESMHPLE 180
 DB 121 ELVDSSNGNFVLELDFEPFNSSPFRPTLSKISGVEFLNRHLSAKLFHDKESMHPLE 180
 QY 181 FLRVHCHGKMMMLNDRIQNLALOHVLRKABEYIGTLPPETPCAEFHRFOELERGW 240
 DB 181 FLRVHCHGKMMMLNDRIQNLALOHVLRKABEYIGTLPPETPCAEFHRFOELERGW 240
 QY 241 GDTARVLEMIQLLDLLEATDPTCEKFLRIPWVENVVLTLPHGVAQDNVGLPPTG 300
 DB 241 GDTARVLEMIQLLDLLEATDPTCEKFLRIPWVENVVLTLPHGVAQDNVGLPPTG 300
 QY 241 GDSARVLESIQLLDLLEATDPTCEKFLRIPWVENVVLTLPHGVAQDNVGLPPTG 300
 DB 241 GDSARVLESIQLLDLLEATDPTCEKFLRIPWVENVVLTLPHGVAQDNVGLPPTG 300
 QY 301 GQVVYILDOVBALENEMLRIRKQGLNTTPRILITRLLPDAVGTTCQRLKRYGTEHS 360
 DB 301 GQVVYILDOVBALENEMLRIRKQGLNTTPRILITRLLPDAVGTTCQRLKRYGTEHS 360
 QY 361 DILRVPFTEKGIYRKWISREKRWPYETTEDVAHEISKELHGTPLLIGNSDGNIV 420
 DB 361 DILRVPFTEKGIYRKWISREKRWPYETTEDVAHEISKELHGTPLLIGNSDGNIV 420
 QY 361 HILRVPFDOKGIYRKWISREKRWPYETTEDVAHEISKELHGTPLLIGNSDGNIV 419
 DB 361 HILRVPFDOKGIYRKWISREKRWPYETTEDVAHEISKELHGTPLLIGNSDGNIV 419
 QY 421 ASLNAHKLGVCTCTTAHALEKTKYPSDSIYWKKEBDKHFSCQFTADLFPANNHTDEIITS 480
 DB 420 ASLNAHKLGVCTCTTAHALEKTKYPSDSIYWKKEBDKHFSCQFTADLFPANNHTDEIITS 479
 QY 481 TFOEIASGSDTVGOYESHTAFTLPGLYRVVHGIDVDFPKFNIYVSGADMEIYFPYTEER 540
 DB 480 TFOEIASGSDTVGOYESHTAFTLPGLYRVVHGIDVDFPKFNIYVSGADMEIYFPYTEER 539
 QY 541 RLKHPHEIEDLLYTKVNEEHLVYNDNRKPIILFMPRLDVKVNLTLGVYCGNPKYR 600
 DB 540 RLKHPHEIEDLLYTKVNEEHLVYNDNRKPIILFMPRLDVKVNLTLGVYCGNPKYR 599
 QY 601 ELANLVVGGRRRKSKELEKAEKMKFELIDKYNLNGQFWMISSQNNRINVELYRYI 660
 DB 600 ELVNLVVAAGRRRKSKELEKAEKMKFELIDKYNLNGQFWMISSQNNRINVELYRYI 659
 QY 661 CDTKCAFQOPALYEAFTLVVAMTCGLPTPATCNGPAEIIYHGKSGFNIDPYHGDAA 720
 DB 660 CDTKCAFQOPAYEAFGLTVVAMTCGLPTPATCNGPAEIIYHGKSGFNIDPYHGDAA 719
 QY 721 DILVDFEKKCKDPBHMWKISQGLKRIEKKYTWKYSERLLTLTGYYGFMKGVNLEER 780
 DB 720 DILVDFEKKCKDPBHMWKISQGLKRIEKKYTWKYSERLLTLTGYYGFMKGVNLEER 779
 QY 781 ESRRLYEMFYALKYRKLAESVPLAE 806
 DB 780 ESRRLYEMFYALKYRKLAESVPLAE 805

RESULT 4

RA Rameberger U., Wedler H., Balke K., Wedler B., Peters S.,
 RA van Stevenen M., Dickse W., Moolman P., Klein Lankhorst R.,
 RA Weltzenegger T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villarroel J., Gielens J., Ardiles W.,
 RA Bents O., Lemke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
 RA Scheller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.,
 RT "Sequence and analysis of chromosome 5 of the plant *Arabidopsis thaliana*".
 RT Nature 408:823-826(2000).
 RL Nature 408:823-826(2000).
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 CC fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
 CC sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X70990; CAA50317.1; -
 DR EMBL: AF296832; -; NOT ANNOTATED CDS.
 DR InterPro: IPR001296; Glyco_trans_1.
 DR InterPro: IPR000368; Sucrose_synth.
 DR Pfam: PF00534; Glycosyl_transf_1; 1.
 DR Pfam: PF00862; Sucrose_synth; 1.
 DR Transferase: Glycosyltransferase; Multigene family.
 KM TRANSFERASE; Glycosyltransferase; Multigene family.
 FT CONFLICT 61 61 R -> Q (IN REF. 1).
 FT CONFLICT 108 108 V -> L (IN REF. 1).
 FT CONFLICT 222 222 S -> P (IN REF. 1).
 FT CONFLICT 405 405 N -> D (IN REF. 1).
 FT CONFLICT 434 434 C -> Q (IN REF. 1).
 FT CONFLICT 609 609 V -> I (IN REF. 1).
 FT CONFLICT 751 751 EK -> DE (IN REF. 1).
 FT CONFLICT 751 751 EK -> DE (IN REF. 1).
 SQ SEQUENCE 806 AA; 92798 MW; 201D365720D46A3F CRC64;

Query Match 84.5%; Score 3569.5; DB 1; Length 806;
 Best Local Similarity 83.9%; Pred. No. 1.4e-223;
 Matches 675; Conservative 66; Mismatches 61; Indels 3; Gaps 2;

QY 2 AERATRVASLRERDETLAHRNFIALLSRIEGKGIIOHNOIIEFALPBRK 61
 DB 4 AERMTTRVHSGRERNETLVSRNEVALLSLRVEAKGKIIOQNOIIEFALPQTRK 63
 QY 62 LANGAFVEYLKASQEAIVLPWVALAVRPRGWEYIRVNVVALVEELTVAEYLHFKEE 121
 DB 64 LEGGFDDLSKSTQERIVLPWVALAVRPRGWEYIRVNVVALVEELTVAEYLHFKEE 123
 QY 122 LVDGSSNGNFVLELDFEPNSSFPRLTSLKSIQNGVEFLNRHLSAKLPDKESMPLLEF 181
 DB 124 LVDGKNGNFVLELDFEPNNAISIPRLTHKYIGNGVDFNRHLSAKLPDKESLPLLEF 183
 QY 182 LRVHKGKGMNMLNDRIQVNLALQVLRKAEVYLGTPETCAEFERFQSIGERMG 241
 DB 184 LRLSHQGNLMLSEKIQVNLALQVLRKAEVYLGTPETCAEFERFQSIGERMG 243
 QY 242 DRAEVLVLEMIQVNLALQVLRKAEVYLGTPETCAEFERFQSIGERMG 301
 DB 244 DRAEVLVLEMIQVNLALQVLRKAEVYLGTPETCAEFERFQSIGERMG 303
 QY 302 QVVYVLDQVRALENEMLRIRKQGNITPRILITRLPDAVGTTCGRLEKYGTSHSD 361
 DB 304 QVVYVLDQVRALENEMLRIRKQGNITPRILITRLPDAVGTTCGRLEKYGTSHSD 363
 QY 362 ILRVPRTEKGIIVRKWISFEKVPWPLETYTDVAHSEIKELHGPDLITGXSGNIVA 421
 DB 364 ILRVPRTEKGIIVRKWISFEKVPWPLETYTDVAHSEIKELHGPDLITGXSGNIVA 422
 QY 422 SLAAKLGVTOCTIAHALEKTKYPPSDIYWKLEDEYHSCQPTADLFAMNHTDPIITST 481

DB 423 SLAAKLGVTOCTIAHALEKTKYPPSDIYWKLEDEYHSCQPTADLFAMNHTDPIITST 482
 QY 482 FOETAGSKDTGVQVESHAFITLPGIYRVHGIIDVDFPENIVSPADMEIYFPYEEKR 541
 DB 483 FOETAGSKDTGVQVESHAFITLPGIYRVHGIIDVDFPENIVSPADMEIYFPYEEKR 542
 QY 542 LKHFEPIEDLLTYTVEENHCVLNDNRKILFTMPRLDRVKULITGLVWCGKPKLRE 601
 DB 543 LKHFEPIEDLLTYTVEENHCVLNDNRKILFTMPRLDRVKULITGLVWCGKPKLRE 602
 QY 602 LANLVVGGDRKESKDLLEKAEKMKFELIDKYNLNGQFWMISQNMIRNVELRYIC 661
 DB 603 LANLVVGGDRKESKDLLEKAEKMKFELIDKYNLNGQFWMISQNMIRNVELRYIC 662
 QY 662 DTKGAFVOPALYEAAGLTVEAMTGLTTPATCNGSPREITVHSGSGNIDPYNDOAD 721
 DB 663 DTKGAFVOPALYEAAGLTVEAMTGLTTPATCNGSPREITVHSGSGNIDPYNDOAD 722
 QY 722 ILVDFEKKDPSHMDKISQGLKRIEKTWKYISRLITLTVSGFMKVSNLRE 781
 DB 723 TLADFTCKEDPSHMDKISQGLKRIEKTWKYISRLITLTVSGFMKVSNLRE 782
 QY 782 SRRYLEMFYALKYRKLAESVPLAER 806
 DB 783 SRRYLEMFYALKYRKLAESVPLAER 805

RESULT 6
 SUSY_ALNGL STANDARD; PRT; 803 AA.
 ID SUSY_ALNGL
 AC P49034;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase).
 GN SUS1.
 OS *Alnus glutinosa* (Alder).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Euxosida I; Fagales; Betulaceae; *Alnus*.
 OX NCBI_TaxID=3517;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Root;
 RX MEDLINE=96186875; PubMed=8602161;
 RA van Gheule M., Pawlowski K.,
 RA Biseling T., Pawlowski K.,
 RT "Sucrose synthase: comparison with legume nodules".
 RL Mol. Gen. Genet. 250:437-446(1996).
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 CC fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
 CC sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X92378; CAA61322.1; -
 DR InterPro: IPR001296; Glyco_trans_1.
 DR InterPro: IPR000368; Sucrose_synth.
 DR Pfam: PF00534; Glycosyl_transf_1; 1.
 DR Pfam: PF00862; Sucrose_synth; 1.
 DR Transferase: Glycosyltransferase.
 SQ SEQUENCE 803 AA; 91630 MW; 28C23736DDF1C731 CRC64;

Query Match 82.8%; Score 3515.5; DB 1; Length 803;

Best Local Similarity 83.8%; Pred. No. 8,1e-219; Indels 5; Gaps 5;
Matches 676; Conservative 56; Mismatches 70;

QY 1 MAERALTIVHSRLRELDLTLAHRNEITLALSRIGKSGKGLIOHQIILFEALPEERK 60
1 MAERLTVHSHRRLDELTVANRHEIYALLSRIGKSGKGLIOHQIILFEALPEERK 60
DB 1 MAERLTVHSHRRLDELTVANRHEIYALLSRIGKSGKGLIOHQIILFEALPEERK 60
QY 61 KLANGAFEEVLKASQGEAIVLPWPVALAVRPRGWYIRVNVNVALVVEELTVAEYLHKE 120
61 KLANGAFEEVLKASQGEAIVLPWPVALAVRPRGWYIRVNVNVALVVEELTVAEYLHKE 120
DB 61 KLANGAFEEVLKASQGEAIVLPWPVALAVRPRGWYIRVNVNVALVVEELTVAEYLHKE 120
QY 121 ELVDSGSGNFVLELDFEPFNASSFPRLTSLKSGNGVEFLNRHLSAKLPHKESMHPLE 180
121 ELVDSGSGNFVLELDFEPFNASSFPRLTSLKSGNGVEFLNRHLSAKLPHKESMHPLE 180
DB 121 ELVDSGSGNFVLELDFEPFNASSFPRLTSLKSGNGVEFLNRHLSAKLPHKESMHPLE 180
QY 181 FLRVCHGKGMMLNDRIQNLNALQHVLRKAEYLGTPPETPCAEFFHFOEIGLEKRGW 240
181 FLRVCHGKGMMLNDRIQNLNALQHVLRKAEYLGTPPETPCAEFFHFOEIGLEKRGW 240
DB 181 FLRVCHGKGMMLNDRIQNLNALQHVLRKAEYLGTPPETPCAEFFHFOEIGLEKRGW 240
QY 241 GDTAERVLEMIQLLDLLEATDPTLKEFLGRIPMFVFNVLITPHGYFAODNVLGYPDTG 300
241 GDTAERVLEMIQLLDLLEATDPTLKEFLGRIPMFVFNVLITPHGYFAODNVLGYPDTG 300
DB 241 GDTAERVLEMIQLLDLLEATDPTLKEFLGRIPMFVFNVLITPHGYFAODNVLGYPDTG 300
QY 301 GQVVTILQVRLNEMMLRIKQOGLNTPRLITLITRLLPDVAGTTCGRLKRYGTSHS 360
301 GQVVTILQVRLNEMMLRIKQOGLNTPRLITLITRLLPDVAGTTCGRLKRYGTSHS 360
DB 301 GQVVTILQVRLNEMMLRIKQOGLNTPRLITLITRLLPDVAGTTCGRLKRYGTSHS 360
QY 361 DILRVPRTKGIYKWIISREFKWPVYETVEDVAHEISKEHGTPLDILGNXSDGIV 420
361 DILRVPRTKGIYKWIISREFKWPVYETVEDVAHEISKEHGTPLDILGNXSDGIV 420
DB 361 DILRVPRTKGIYKWIISREFKWPVYETVEDVAHEISKEHGTPLDILGNXSDGIV 420
QY 421 ASLNAKLVGTQCTIAHALKTKYPSDSIYWKQKLBKXFSQCTADLFAAMHTDFIITS 480
421 ASLNAKLVGTQCTIAHALKTKYPSDSIYWKQKLBKXFSQCTADLFAAMHTDFIITS 480
DB 421 ASLNAKLVGTQCTIAHALKTKYPSDSIYWKQKLBKXFSQCTADLFAAMHTDFIITS 480
QY 481 TFOEIASGKDTYGOESTHTAFLPGLYRVHAGIDVDFDKPNVVSAGAMEIYFPTTEKR 540
481 TFOEIASGKDTYGOESTHTAFLPGLYRVHAGIDVDFDKPNVVSAGAMEIYFPTTEKR 540
DB 481 TFOEIASGKDTYGOESTHTAFLPGLYRVHAGIDVDFDKPNVVSAGAMEIYFPTTEKR 540
QY 541 RLKHHPRLEDDLTYKVENEBELCYLNDNRKPIELFTMPRLDVKVLITGLVEMCGNPKLR 600
541 RLKHHPRLEDDLTYKVENEBELCYLNDNRKPIELFTMPRLDVKVLITGLVEMCGNPKLR 600
DB 541 RLKHHPRLEDDLTYKVENEBELCYLNDNRKPIELFTMPRLDVKVLITGLVEMCGNPKLR 600
QY 597 ELVNLVVAAGNLEKSKDNEBEKAMTKHGLIETKYNLQGFWRMISSQMMRVANGELRYXI 656
597 ELVNLVVAAGNLEKSKDNEBEKAMTKHGLIETKYNLQGFWRMISSQMMRVANGELRYXI 656
DB 597 ELVNLVVAAGNLEKSKDNEBEKAMTKHGLIETKYNLQGFWRMISSQMMRVANGELRYXI 656
QY 601 ELANLVVGGDRKESKDLBEKAKEMKMFELIDKYNLNGOFWRMISSQMMRIANVELRYXI 660
601 ELANLVVGGDRKESKDLBEKAKEMKMFELIDKYNLNGOFWRMISSQMMRIANVELRYXI 660
DB 601 ELANLVVGGDRKESKDLBEKAKEMKMFELIDKYNLNGOFWRMISSQMMRIANVELRYXI 660
QY 661 CDTKAP-VQPALYFAFGITVVEAMTCGIPFATCNGSPALITVYKSGGNIDPYHGOA 719
661 CDTKAP-VQPALYFAFGITVVEAMTCGIPFATCNGSPALITVYKSGGNIDPYHGOA 719
DB 661 CDTKAP-VQPALYFAFGITVVEAMTCGIPFATCNGSPALITVYKSGGNIDPYHGOA 719
QY 720 ADILVDFEKKCKDSDHMDKISQGLKRIEKEYTWKIVSERLLTLTGYYGFWKHSNLER 779
720 ADILVDFEKKCKDSDHMDKISQGLKRIEKEYTWKIVSERLLTLTGYYGFWKHSNLER 779
DB 720 ADILVDFEKKCKDSDHMDKISQGLKRIEKEYTWKIVSERLLTLTGYYGFWKHSNLER 779
QY 777 LESRRYIEMFVALKYRKLAESVPLAVE 803
777 LESRRYIEMFVALKYRKLAESVPLAVE 803
DB 777 LESRRYIEMFVALKYRKLAESVPLAVE 803

QX NCBI_TaxID=4113;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Flab67; TISSUE=leaf;
RX MEDLINE=88152501; PubMed=2964386;
RA Salanoubat M., Belliard G.;
RT "Molecular cloning and sequencing of sucrose synthase cDNA from
potato (Solanum tuberosum L.): preliminary characterization of
sucrose synthase mRNA distribution.";
RL Gene 60:47-56 (1987).
RM [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Flab67; TISSUE=leaf;
RX MEDLINE=96172787; PubMed=589622;
RA Fu H., Park W.D.;
RT "Sink- and vascular-associated sucrose synthase functions are encoded
by different gene classes in potato.";
RL Plant Cell 7:1369-1385 (1995).
CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
fructose for various metabolic pathways.
CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
CC -1- TISSUE SPECIFICITY: Expression is at least 10 fold higher in
tubers compared to photosynthetically active tissues.
CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
sucrose synthase subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by, and for commercial
entities requires a license agreement (See <http://www.1db-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M18745; AAA33841.1; -;
DR EMBL: U24087; AAA97571.1; -;
DR PIR: A29615; YUFOS.
DR InterPro: IPR001296; Glyco_trans_1.
DR InterPro: IPR000368; Sucrose_synth.
DR Pfam: PF00534; Glycos_transf_1.
DR Pfam: PF00862; Sucrose_synth_1.
KM Transferrase, Glycosyltransferase, Multigene family.
FT CONFLICT 221 221 E -> D (IN REF. 2).
FT CONFLICT 731 732 K D -> R (IN REF. 2).
FT CONFLICT 741 741 M -> T (IN REF. 2).
FT CONFLICT 748 748 E -> Q (IN REF. 2).
FT CONFLICT 759 759 S -> R (IN REF. 2).
SQ SEQUENCE 805 AA; 92416 MW; C453363A7CD32809 CRC64;
Query Match 82.4%; Score 3501.5; DB 1; Length 805;
Best Local Similarity 82.1%; Pred. No. 6,5e-218;
Matches 662; Conservative 65; Mismatches 78; Indels 1; Gaps 1;

QY 1 MAERALTIVHSRLRELDLTLAHRNEITLALSRIGKSGKGLIOHQIILFEALPEERK 60
1 MAERLTVHSHRRLDELTVANRHEIYALLSRIGKSGKGLIOHQIILFEALPEERK 60
DB 1 MAERLTVHSHRRLDELTVANRHEIYALLSRIGKSGKGLIOHQIILFEALPEERK 60
QY 61 KLANGAFEEVLKASQGEAIVLPWPVALAVRPRGWYIRVNVNVALVVEELTVAEYLHKE 120
61 KLANGAFEEVLKASQGEAIVLPWPVALAVRPRGWYIRVNVNVALVVEELTVAEYLHKE 120
DB 61 KLANGAFEEVLKASQGEAIVLPWPVALAVRPRGWYIRVNVNVALVVEELTVAEYLHKE 120
QY 121 ELVDSGSGNFVLELDFEPFNASSFPRLTSLKSGNGVEFLNRHLSAKLPHKESMHPLE 180
121 ELVDSGSGNFVLELDFEPFNASSFPRLTSLKSGNGVEFLNRHLSAKLPHKESMHPLE 180
DB 121 ELVDSGSGNFVLELDFEPFNASSFPRLTSLKSGNGVEFLNRHLSAKLPHKESMHPLE 180
QY 181 FLRVCHGKGMMLNDRIQNLNALQHVLRKAEYLGTPPETPCAEFFHFOEIGLEKRGW 240
181 FLRVCHGKGMMLNDRIQNLNALQHVLRKAEYLGTPPETPCAEFFHFOEIGLEKRGW 240
DB 181 FLRVCHGKGMMLNDRIQNLNALQHVLRKAEYLGTPPETPCAEFFHFOEIGLEKRGW 240
QY 241 GDTAERVLEMIQLLDLLEATDPTLKEFLGRIPMFVFNVLITPHGYFAODNVLGYPDTG 300
241 GDTAERVLEMIQLLDLLEATDPTLKEFLGRIPMFVFNVLITPHGYFAODNVLGYPDTG 300
DB 241 GDTAERVLEMIQLLDLLEATDPTLKEFLGRIPMFVFNVLITPHGYFAODNVLGYPDTG 300

```

QY 301 GQVYVILDOVALRENMELRLIKOQGLNTPTRLITRLRLPAVGTGCGRLKXYGTSHS 360
D 301 GQVYVILDOVALRENMELRLIKOQGLNTPTRLITRLRLPAVGTGCGRLKXYGTSHS 360
QY 361 DILRVPTTEKGIYRKWISREPKWPLYETTEDVAHEISKEHGTPDLIIGNNSDGNIV 420
D 361 HILRVPTTEKGIYRKWISREPKWPLYETTEDVAHEISKEHGTPDLIIGNNSDGNIV 420
QY 421 ASLAAHKLGVTQCTTAHAEKTKYPSDSIYWKKEDEKXHFSCQFTADLIANNHTDFIITS 480
D 421 ASLAAHKLGVTQCTTAHAEKTKYPSDSIYWKKEDEKXHFSCQFTADLIANNHTDFIITS 480
QY 481 TFOEIASGKDVGVGYESHMTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEERK 540
D 481 TFOEIASGKDVGVGYESHMTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEERK 540
QY 541 RLKHFPEIEDLLTYKVENEBHLCVLNDRNKPIFTMPRLDRVKNLTGLVEMCGKNPRLR 600
D 541 RLKHFPEIEDLLTYKVENEBHLCVLNDRNKPIFTMPRLDRVKNLTGLVEMCGKNPRLR 600
QY 601 ELANLVVVGGRRKESKOLEEAKEMKMFELIDKYNLNGQPRWISSQNNRIRANVELYRYI 660
D 601 ELANLVVVGGRRKESKOLEEAKEMKMFELIDKYNLNGQPRWISSQNNRIRANVELYRYI 660
QY 661 CDTGAFVQPAALYEAFTLVVEAMTCGLPTPATCNGPABEIIYHGKSGFNIDPYHGDOAA 720
D 661 ADTKGAFVQPAALYEAFTLVVEAMTCGLPTPATCNGPABEIIYHGKSGFNIDPYHGDOAA 720
QY 721 DILVDFEKKCKDPSHMDKISOGGLKRIEKKYTKYISERLLTLTGYYGFWMKHSNLERR 780
D 721 DILVDFEKKCKDPSHMDKISOGGLKRIEKKYTKYISERLLTLTGYYGFWMKHSNLERR 780
QY 781 ESRRLMEFYALKYRKLAESVPLAE 806
D 781 ESRRLMEFYALKYRKLAESVPLAE 806
QY 780 EIRRYLEMFYALKYRKMAEAVPLAE 805
D 780 EIRRYLEMFYALKYRKMAEAVPLAE 805

```

RESULT 8

SUSY_LYCES
ID SUSY_LYCES STANDARD; PRT; 805 AA.

```

AC P49037;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP-glucosyltransferase).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. VF36; TISSUE=Placil;
RX MEDLINE=94120019; PubMed=8290642;
RA Wang F., Smith A.G., Brenner M.L.;
RT "Isolation and sequencing of tomato fruit sucrose synthase cDNA.";
RL Plant Physiol. 103:1463-1464 (1993).
CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
CC fructose for various metabolic pathways.
CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
CC sucrose synthase subfamily.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; L19762; AAA34196.1; -
 DR InterPro; IPR001296; Glyco_trans_1.

DR InterPro; IPR000368; Sucrose synth.
 DR Pfam; PF00534; Glycos_transf_1.
 DR Pfam; PF00862; Sucrose_synth_1.
 KM Transferase; Glycosyltransferase.
 SQ SEQUENCE 805 AA; 92446 MW; 8C7CCB09415483F5 CRC64;

Query Match 82.3%; Score 3496.5; DB 1; Length 805;
 Best Local Similarity 81.9%; Pred. No. 1.4e-217;
 Matches 660; Conservative 67; Mismatches 78; Indels 1; Gaps 1;

```

QY 1 MAERLTVHSLRERDELTLAHRNEILLALISIEGKGILQHHQIILEFAIPEENRK 60
D 1 MAERLTVHSLRERDELTLAHRNEILLALISIEGKGILQHHQIILEFAIPEENRK 60
QY 61 KLANGAFPEVLAQOEAIVLPWVALAVRPFGWEYIRRVNVALVYBELTVAEYIHRKE 120
D 61 KLANGAFPEVLAQOEAIVLPWVALAVRPFGWEYIRRVNVALVYBELTVAEYIHRKE 120
QY 121 ELVDSGSGNFVLEIDPEFPFSPPRLTSSKISNGVYEFNRRHSAXLFDHESMPLLE 180
D 121 ELVDSGSGNFVLEIDPEFPFSPPRLTSSKISNGVYEFNRRHSAXLFDHESMPLLE 180
QY 181 FLRVYCHGKNMMDNRIQNLNALQHVLRKAEYLGTLPEPTCAEFERHPOEIGLERGW 240
D 181 FLRVYCHGKNMMDNRIQNLNALQHVLRKAEYLGTLPEPTCAEFERHPOEIGLERGW 240
QY 241 GDTAEVLEWCMILDLLEAPDSCTLEKFLORIMVFNVVILSPHGFAQENVAGYPTDG 300
D 241 GDTAEVLEWCMILDLLEAPDSCTLEKFLORIMVFNVVILSPHGFAQENVAGYPTDG 300
QY 301 GQVYVILDOVALRENMELRLIKOQGLNTPTRLITRLRLPAVGTGCGRLKXYGTSHS 360
D 301 GQVYVILDOVALRENMELRLIKOQGLNTPTRLITRLRLPAVGTGCGRLKXYGTSHS 360
QY 361 DILRVPTTEKGIYRKWISREPKWPLYETTEDVAHEISKEHGTPDLIIGNNSDGNIV 420
D 361 HILRVPTTEKGIYRKWISREPKWPLYETTEDVAHEISKEHGTPDLIIGNNSDGNIV 420
QY 421 ASLAAHKLGVTQCTTAHAEKTKYPSDSIYWKKEDEKXHFSCQFTADLIANNHTDFIITS 480
D 421 ASLAAHKLGVTQCTTAHAEKTKYPSDSIYWKKEDEKXHFSCQFTADLIANNHTDFIITS 480
QY 481 TFOEIASGKDVGVGYESHMTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEERK 540
D 481 TFOEIASGKDVGVGYESHMTAFTLPGLYRVVHGIDVDPKFNIVSPGADMEIYFPYTEERK 540
QY 541 RLKHFPEIEDLLTYKVENEBHLCVLNDRNKPIFTMPRLDRVKNLTGLVEMCGKNPRLR 600
D 541 RLKHFPEIEDLLTYKVENEBHLCVLNDRNKPIFTMPRLDRVKNLTGLVEMCGKNPRLR 600
QY 600 GLVNLVVVGGRRKESKOLEEAKEMKMFELIDKYNLNGQPRWISSQNNRIRANVELYRYI 659
D 600 GLVNLVVVGGRRKESKOLEEAKEMKMFELIDKYNLNGQPRWISSQNNRIRANVELYRYI 659
QY 661 CDTGAFVQPAALYEAFTLVVEAMTCGLPTPATCNGPABEIIYHGKSGFNIDPYHGDOAA 720
D 661 ADTKGAFVQPAALYEAFTLVVEAMTCGLPTPATCNGPABEIIYHGKSGFNIDPYHGDOAA 720
QY 721 DILVDFEKKCKDPSHMDKISOGGLKRIEKKYTKYISERLLTLTGYYGFWMKHSNLERR 780
D 721 DILVDFEKKCKDPSHMDKISOGGLKRIEKKYTKYISERLLTLTGYYGFWMKHSNLERR 780
QY 780 EIRRYLEMFYALKYRKMAEAVPLAE 805
D 780 EIRRYLEMFYALKYRKMAEAVPLAE 805

```

RESULT 9

SUS2_SOLTU
ID SUS2_SOLTU STANDARD; PRT; 805 AA.

AC P49037;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update).

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sucrose synthase (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase)
 DE (S865).
 OS Solanum tuberosum (potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Solanales; Solanaceae; Solanum.
 NCBI_TaxID=4113;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. FL1607; TISSUE=leaf;
 RA MEDLINE=96172787; PubMed=8589622;
 RX Fu H., Park W.D.;
 RT "Sink- and vascular-associated sucrose synthase functions are encoded
 by different gene classes in potato.";
 RL Plant Cell 7:1369-1385(1995).
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glucosyltransferase family 1. Plant
 sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U24088; AAA97572.1; -
 DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR000368; Sucrose synth.
 DR Pfam; PF00534; Glycosyltransferase_1.
 DR Pfam; PF00862; Sucrose_synth_1.
 DR Transferase; Glycosyltransferase; Multigene family.
 DR SEQUENCE 805 AA; 92577 MW; 6117DBD2E4DD35DC CRC64;
 SQ
 Query Match 81.9%; Score 3476.5; DB 1; Length 805;
 Best local similarity 81.9%; Pred. No. 2,7e-216;
 Matches 660; Conservative 65; Mismatches 80; Indels 1; Gaps 1;
 QY 1 MAERALTVHSLERLDELTLAHRNEIILLSRIEKGKGLIQHIIIEFAIPEENRK 60
 DB 1 MAERLVTVHSLERLDELTLAHRNEIILLSRIEKGKGLIKPQLLAEFESIKHEDK 60
 QY 61 KLANAFBEVYLKASGEATVLPWVALAARPRGVWEIYRVNVAHVVELTVAETLHKE 120
 DB 61 KLNDHAFBEVYKSTGEATVLPWVALAIRLRPGVWEYRVNVAHVVELTVAETLHKE 120
 QY 121 ELVDSSNGNFVLELDPEFPNSPPRLTSKISGVGEFLNRHLAKLPHDKESMPLLE 180
 DB 121 ELVNGTSDNDFVLELDPEFPNSPPRLTSKISGVGEFLNRHLAKLPHDKESMPLLE 180
 QY 181 FLRVHCHGKMMMLNDRIQNLALQHLARKAEVYGLTLPETPCAEPFHRPEQIGLRCW 240
 DB 181 FLRVHCHGKMMMLNDRIQNLALQHLARKAEVYGLTLPETPCAEPFHRPEQIGLRCW 240
 QY 241 GPTARVLEMTOLLDLLEATPRLTEKLGRIWVFNVVILTPRGYPAQDNVGLPRTG 300
 DB 241 GPTARVLEMTOLLDLLEATPRLTEKLGRIWVFNVVILTPRGYPAQDNVGLPRTG 300
 QY 301 GGVVILDOVRLALEMELIRIKOGANTPRILITRLLPDAVGTTCORLEKVGSTHS 360
 DB 301 GGVVILDOVRLALEMELIRIKOGANTPRILITRLLPDAVGTTCORLEKVGSTHS 360
 QY 361 DILRVFPTEKGIIVKWSIRFEKWPVYLETYEDVAHEISKELHSTPDLLIGNSDGNIV 420
 DB 361 DILRVFPTEKGIIVKWSIRFEKWPVYLETYEDVAHEISKELHSTPDLLIGNSDGNIV 420
 QY 421 ASLNAHKGVTOCTAAHLEKTKYPSDLYMKLEDKTHFSQCFADLFANMHTDFTTS 480
 DB 421 ASLNAHKGVTOCTAAHLEKTKYPSDLYMKLEDKTHFSQCFADLFANMHTDFTTS 480
 QY 480 TFOEIASKDTVGYGESHRTAFTLPLGLYRVVHGIDVDPKFNIVSPGADMEIYPTTEKR 540
 DB 480 TFOEIASKDTVGYGESHRTAFTLPLGLYRVVHGIDVDPKFNIVSPGADMEIYPTTEKR 540
 QY 541 RLKHFHEIDLTLYTKNEENHLCVNDNRKPLIFTPRILDRVKNLGLVEMCKNKLK 600
 DB 541 RLTFHEIIDLTFSDVNEENHLCVNDNRKPLIFTPRILDRVKNLGLVEMCKNKLK 600
 QY 601 ELANLVVGGDRKRESKDLSEKAEKMFELIDKYNNGOFWMISSQNRIRNVELRYI 660
 DB 601 ELANLVVGGDRKRESKDLSEKAEKMFELIDKYNNGOFWMISSQNRIRNVELRYI 660
 QY 661 CDTKGAFOPALYEAFGLYVEAMTCGLPTPATCNGSPAEIIVHGKSGFNIDPYHQDA 720
 DB 660 ADTGAFVOPAFYBAFGLYVEAMSCGLPTPATCNGSPAEIIVHGKSGFQIDPYHQDA 720
 QY 721 DILVDFPEKCKDSDHMDKISOGGLKRIEKKYTKIYSERLTLTYGFGFKHNSNERR 780
 DB 721 DILVDFPEKCKDSDHMDKISOGGLKRIEKKYTKIYSERLTLTYGFGFKHNSNERR 780
 QY 780 ESRRYLEMFYALKRYKLAESVPLAE 806
 DB 780 ESRRYLEMFYALKRYKLAESVPLAE 806
 QY 806 ESRRYLEMFYALKRYKLAESVPLAE 806
 DB 806 ESRRYLEMFYALKRYKLAESVPLAE 806

RESULT 10
 SUSL_DAUCA STANDARD; PRT; 808 AA.
 ID SUSL_DAUCA
 AC P49035;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Sucrose synthase isoform I (EC 2.4.1.13) (Sucrose-UDP
 glucosyltransferase 1) (Suey*Dcl).
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Campanulidales; Apiales; Apiaceae; Apioidae; Scandiceae; Daucinae;
 OC Daucus.
 NCBI_TaxID=4039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nantaise;
 RX MEDLINE=95303982; PubMed=7784526;
 RA SebkoVA V., Unger C., Hardegger M., Sturm A.;
 RT "Biochemical, physiological, and molecular characterization of
 sucrose synthase from Daucus carota.";
 RL Plant Physiol. 108:75-83(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nantaise; TISSUE=leaf;
 RX MEDLINE=99178785; PubMed=10080700;
 RA Sturm A., Lienhard S., Schacht S., Hardegger M.;
 RT "Tissue-specific expression of two genes for sucrose synthase in
 carrot (Daucus carota L.).";
 RL Plant Mol. Biol. 39:349-360(1999).
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- ENZYME REGULATION: Fructose acts as a noncompetitive inhibitor
 with an inhibition constant of 17.2 mM. In contrast, glucose
 inhibits uncompetitively with an inhibition constant of 4.3 mM.
 CC -1- SUBUNIT: Homotrimer.
 CC -1- TISSUE SPECIFICITY: Expressed in stems, in roots at different
 developmental stages, and in flower buds, flowers and maturing
 seeds, with the highest levels in strong utilization sinks for
 sucrose such as growing stems and tap root tips.
 CC -1- SIMILARITY: Belongs to the glucosyltransferase family 1. Plant
 sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X75332; CAA53081.1; -
 DR EMBL: Y16090; CAA53081.1; -
 DR PIR: S37560; S37560
 DR InterPro: IPR001286; Glyco_trans_1.
 DR InterPro: IPR000368; Sucrose_synth.
 DR Pfam: PF00534; Glyco_transf_1; 1.
 DR Pfam: PF00862; Sucrose_synth; 1.
 DR Transference: Glycosyltransferase; Multigene family.
 KM SEQUENCE 808 AA; 92474 MW; 1A40FCBAA4A5425 CRC64;

Query Match 80.1%; Score 3400.5; DB 1; Length 808;
 Best Local Similarity 79.1%; Pred. No. 2.1e-211;
 Matches 639; Conservative 76; Mismatches 90; Indels 3; Gaps 2;

QY 1 MAERALTVRHSREIRDETLAHRNEILALSRIEGKGIIQHIIIEFAIPENRK 60
 DB 1 MEEPLVTRHSRERDSTLANRNEILMFLSRISHGKILPHQLAEYALISKEDKL 60
 QY 61 KL-ANGAFEEVLKASQEAIVLPWVALAVRPPGVMEYIRVNVVALVEELTVAEYLHF 118
 DB 61 KLDDHGNAFAEVIKQGEAIVSPWVALAIRPGWMEYIRVNVVALVEELTVAEYLHF 120
 QY 119 KEELVDGSSNGNVLDELDEPPNSSFPPRLTSSKISNGVEFLNRHLSAKLPHDKESMPL 178
 DB 121 KEELVIGSSDANFVELDEPPTASPPRLTSSKISNGVEFLNRHLSAKLPHDKESMPL 180
 QY 179 LEFLVHCHKGKMMNDRIQNLNLQHLAKAEVLTGLPPTPAEENHFOETGLER 238
 DB 181 LEFLVHCHKGKMMNDRIQNLNLQHLAKAEVLTGLPPTPAEENHFOETGLER 240
 QY 239 GWDGTAERVLMIQLLDLLEATDPTCEKFLGRIPMVENVVILTPHGYFAADNVLYGYPD 298
 DB 241 GWDGTAERVLMIQLLDLLEATDPTCEKFLGRIPMVENVVILTPHGYFAADNVLYGYPD 300
 QY 299 TGGQVYVILDOVRALENEMLRIRKQGLNITRILITLRLPDVAVGTTGGRLERKYGTE 358
 DB 301 TGGQVYVILDOVRALEREMIKRIKEQGLDIKPRILITLRLPDVAVGTTGGRLERKYGTE 360
 QY 359 HSDILRVPRTEKGIYRKWISRFEEKWPLYETTEDVAHEISKEHLGTDLIIGNXSDGN 418
 DB 361 HAHILRVPRTEKGIYRKWISRFEEKWPLYETTEDVAHEISKEHLGTDLIIGNXSDGN 419
 QY 419 IYASLLAHGLVTOCTIAHLEKTKYPSDSIYWKLEDEKYHSCOPTADLFAMNHTDFII 478
 DB 420 IYASLLAHGLVTOCTIAHLEKTKYPSDSIYWKLEDEKYHSCOPTADLFAMNHTDFII 479
 QY 479 TSTFOEIASGKDTVOQYESHTAFTLPGLYRVVHGIDVDPKKNIVSPGADMEIYEPYTEE 538
 DB 480 TSTFOEIASGKDTVOQYESHTAFTLPGLYRVVHGIDVDPKKNIVSPGADMEIYEPYTEE 539
 QY 539 KRRLGHFHEIEDLTYKNEENHCVLNDKRNKPLTFMPRLDRYKNTGLVWKGKKNPK 598
 DB 540 KRRLGHFHEIEDLTYKNEENHCVLNDKRNKPLTFMPRLDRYKNTGLVWKGKKNPK 599
 QY 599 LRELNLVVGDRRKESKDLSEKAKMKMPELIDKYNLNGOFRMISOMNIRNVLEYLR 658
 DB 600 LRELNLVVGDRRKESKDLSEKAKMKMPELIDKYNLNGOFRMISOMNIRNVLEYLR 659
 QY 659 YICDTKGAFFVOPALYFAFGLTVEAMTCGLPTFATCNGGPAAEIVHGSGFNIDPYHQD 718
 DB 660 YICDTKGAFFVOPALYFAFGLTVEAMTCGLPTFATCNGGPAAEIVHGSGFNIDPYHQD 719
 QY 719 AADILVDFPEKCKDPSHWDKISOGGLKRIEKKYMKIYSEBLLTLTGYYGKWKVSNLE 778
 DB 720 AADILVDFPEKCKDPSHWDKISOGGLKRIEKKYMKIYSEBLLTLTGYYGKWKVSNLE 779
 QY 779 RRESRRYLEMFALYKRYKLAESVPLAE 806

DB 780 RLEIRRYLEMFALYKRYKLAESVPLAKD 807

RESULT 11

SUS2_TULGE STANDARD; PRT; 820 AA.
 ID SUS2_TULGE
 AC 041607;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT Sucrose synthase 2 (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase 2).
 OS Tulipa gesneriana (Tulip).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
 OC Tulipa.
 NC NCBITaxID=13306;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Apeldoorn;
 RA Balk P.A., de Boer A.D.;
 RL Submitted (Mar-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 CC fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
 CC sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: X96939; CAA56440.1; -

InterPro: IPR001296; Glyco_trans_1.

InterPro: IPR000368; Sucrose_synth.

Pfam: PF00534; Glyco_transf_1; 1.

Pfam: PF00862; Sucrose_synth; 1.

Transference: Glycosyltransferase; Multigene family.

SEQUENCE 820 AA; 93015 MW; AA75AF8BD74A0BED CRC64;

Query Match 77.7%; Score 3301; DB 1; Length 820;
 Best Local Similarity 77.0%; Pred. No. 5.5e-205;
 Matches 620; Conservative 88; Mismatches 95; Indels 2; Gaps 2;

QY 1 MAERALTVRHSREIRDETLAHRNEILALSRIEGKGIIQHIIIEFAIPENRK 59
 DB 1 MADRMNTRVHSRERDSTLANRNEILMFLSRISHGKILPHQLAEYALISKEDKL 60
 QY 60 KKLANGAFEEVLKASQEAIVLPWVALAVRPPGVMEYIRVNVVALVEELTVAEYLHF 119
 DB 61 KKLANGAFEEVLKASQEAIVLPWVALAVRPPGVMEYIRVNVVALVEELTVAEYLHF 120
 QY 120 BELVUGSSNGNVLDELDEPPNSSFPPRLTSSKISNGVEFLNRHLSAKLPHDKESMPL 179
 DB 121 BELVUGSSNGNVLDELDEPPNSSFPPRLTSSKISNGVEFLNRHLSAKLPHDKESMPL 180
 QY 180 EFLRVHCHKGKMMNDRIQNLNLQHLAKAEVLTGLPPTPAEENHFOETGLER 239
 DB 181 EFLRVHCHKGKMMNDRIQNLNLQHLAKAEVLTGLPPTPAEENHFOETGLER 240
 QY 240 WGDGTAERVLMIQLLDLLEATDPTCEKFLGRIPMVENVVILTPHGYFAADNVLYGYPD 299
 DB 241 WGDGTAERVLMIQLLDLLEATDPTCEKFLGRIPMVENVVILTPHGYFAADNVLYGYPD 300
 QY 300 GGOVYVILDOVRALENEMLRIRKQGLNITRILITLRLPDVAVGTTGGRLERKYGTEH 359
 DB 301 GGOVYVILDOVRALENEMLRIRKQGLNITRILITLRLPDVAVGTTGGRLERKYGTEH 360
 QY 360 SDILRVPRTEKGIYRKWISRFEEKWPLYETTEDVAHEISKEHLGTDLIIGNXSDGN 419

Db 726 LTVDFDCKQAEPSWMSKISGGCLORIEKKYTWKLYSERLMTLTGVYGFWMKYNLERRE 785
 Qy 782 SRRYLEMFYALKYRKLAESVPLAEE 806
 Db 786 TRRYLEMYALKYRKMASTVPLAVE 810

RESULT 13
 SUS1_TULGE STANDARD: PRT: 805 AA.

AC Q41608: Rel. 37, Created
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Sucrose synthase 1 (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase 1).
 OS Tulipa gesneriana (Tulip)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
 OC Tulipa.
 OC NCBI_TaxID=13306;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Apeldoorn;
 RA Balk P.A., de Boer A.D.;
 RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
 sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use. By non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, X96938; CAA65639.1; -
 DR InterPro; IPR001296; Glyco_trans_1.
 DR InterPro; IPR000368; Sucrose_synth.
 DR Pfam; PF00534; Glycos_transf_1; 1.
 DR Pfam; PF00862; Sucrose_synth; 1.
 KM Transferrase; Glycosyltransferase; Multigene family.
 SQ SEQUENCE 805 AA; 92254 MW; B749D8953C938F0 CRC64;

Query Match 77.1%; Score 3273; DB 1; Length 805;
 Best Local Similarity 76.1%; Pred. No. 3,4e-203;
 Matches 613; Conservative 97; Mismatches 91; Indels 4; Gaps 3;

Qy 1 MAERALTIVHSGIRELDETLLAHRREITALLSRIGKGGKGLIHOHIIIEBA-IPEENR 59
 Db 1 MTERLTITIHISIKERLGDLSHHHPBELALTSRFRKCKGMLERHQLTLEVESVPEADR 60
 Qy 60 KKLNGAFPEVLKASQEAIVLPWVALAVRPPGWEIVRVVVALVVEELTVAEYLFK 119
 Db 61 EKLKQVEDTLKRSQEAIVLPWVALAIRPPGWEIVRVVVALVVEELTVAEYLFK 118
 Qy 120 BELVDSSNGNFVLELDEPFNPSPPRPTLSKISNGVEFLNRLHSALFLDKESMPL 179
 Db 119 EDVDRSSQSNFVLEMFEPFANVPRLSPSLKISNGVQFLNRLHSALFLDKESLTYLL 178
 Qy 180 EFLRVHCHKKGMINDRIQNLALQHVLRKAEEYLGTPPETPAEERHFOETGLERG 239
 Db 179 NPLFHNHVKGTLLMINDRLQSLSAQTLARKADRYLLSISKDTPYSEFNHSGFOYGLRKG 238
 Qy 240 WGDFAERYLEMIQLLDLEATDPCTLEKFLGRIMVENVVITLPHGYFAODNVLYGPD 299
 Db 239 WGDTRSRSENIHLDDLLEAPDPTLEKFLGTTIMVENVVILSHGFAQANVLYGPD 298
 Qy 300 GGOVVYILDQVRALENEMLRKIQGLNITPRILITLPLDPAVGTGGRLEKRYGTEH 359

Db 229 GGOVVYILDQVRALETEMLKIKOQGLDITPRILITLPLDPAVGTGGRLEKRYGTEH 358
 Qy 360 SDILRVPRTEKGIYRKWISREKWPYLETYTEDEVAHEISKEHLGTPDLIIKXNSDGI 419
 Db 359 THILRVPRIDGILRKWISRFE-WPYLETYTEDEVAHELGEMOATPDLIIKXNSDGI 417
 Qy 420 VASILAHRLGVTOCTIAHALEKTKYPSDSIYWKLEDKYHSPCOPTADLFANHTDFTIT 479
 Db 418 VASILAHRLGVTOCTIAHALEKTKYPSDSIYWKLEDKYHSPCOPTADLFANHTDFTIT 477
 Qy 480 STFOELIASSKDTVGQYESHTAFTLPGLYRVVHGIDVPKKNIVSPGADMEIYPTTEK 539
 Db 478 STFOELIASSKDTVGQYESHTAFTLPGLYRVVHGIDVPKKNIVSPGADMEIYPTTEK 537
 Qy 540 RLKHFHEIDLLYTKYNEEHCVLDRNKPILFTMPRLDRVYNLGLVEMGSKNKL 599
 Db 538 KRLTRAHAEIELLYSVNEEHKFLVDKRNKPIIFSNARIDRVNMGVLELKGKRL 597
 Qy 600 RELANLVVGGDRRKESKOLEEAKMKAFELIDKYNLNGOFRWISSOMNRIRNVELYR 659
 Db 598 KELVNLVVVADHGKESKOLEEAKMKAFELIDKYNLNGOFRWISSOMNRIRNVELYR 657
 Qy 660 ICTYKAFVQPALYFAFGLVVEAMTCGLPTPATCGGPAEIIYHKGSGFNIDPYHQA 719
 Db 658 IADTGAFVQPAFYEAFFGLTVESWTCGLPTPATCGGPAEIIYHKGSGFNIDPYHQA 717
 Qy 720 ADIYVDFEKKKQPSHDKISGGCLORIEKKYTWKLYSERLMTLTGVYGFWMKYNLER 779
 Db 718 SEQVSPFEKKEKEDPAHWEKISGGCLORIEKKYTWKLYSERLMTLTGVYGFWMKYNLDR 777

780 RESRRYLEMFYALKYRKLAESVPLA 804
 778 TRRYLEMYALKYRKMASTVPLA 802

RESULT 14
 SUS2_ORYSA STANDARD: PRT: 816 AA.

AC P31924:
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sucrose synthase 2 (EC 2.4.1.13) (Sucrose-UDP glucosyltransferase 2).
 GN SS2.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Euphorbiaceae; Oryzae; Oryza.
 OC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Tainong 67;
 RX MEDLINE=92119223; PubMed=1531032;
 RA Yu W.P., Wang A.Y., Huang R.H., Sung H.Y., Su J.C.;
 RT "Isolation and sequences of rice sucrose synthase cDNA and genomic
 DNA";
 RL Plant Mol. Biol. 18:139-142 (1992).
 RN [2]
 RP SEQUENCE OF 605-816 FROM N.A.
 RX MEDLINE=92288314; PubMed=1534703;
 RA Wang A.Y., Yu W.P., Huang R.H., Huang J.W., Sung H.Y., Su J.C.;
 RT "Presence of three rice sucrose synthase genes as revealed by cloning
 and sequencing of cDNA";
 RL Plant Mol. Biol. 18:1191-1194 (1992).
 CC -1- FUNCTION: Sucrose-cleaving enzyme that provides UDP-glucose and
 fructose for various metabolic pathways.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose = UDP + sucrose.
 CC -1- SIMILARITY: Belongs to the glycosyltransferase family 1. Plant
 sucrose synthase subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its

Db 243 RVLDTLHLLDLLEAPDPALEKFLGTIPMMFNVLSPHGYFAQSNVLGYPDITGGQVYV 302
Qy 306 ILDOVRALLENMLARIKQOGLNITPRILITRLLPDAGTTGGRLKRYGTEHSDILRV 365
Db 303 ILDOVRALLENMLARIKQOGLDITPKLITRLLPDAGTTGGRLKRYGTEHSDILRV 362
Qy 366 PPRTEKGIVRKWIISFEKWPYLETYTEDVAHEISKEHGTPDLLIGNXSDGNIVASLLA 425
Db 363 PPRNENGLRKWKISRFD-VMPYLETYTEDVSEIMKEMQAKPDLLIGNYSNGNLVATLLA 421
Qy 426 HKLGVTCTIAHALEKTYPPSDIYWKLEDEKYPSCQFTADLFAMNHTDPIITSTFOEI 485
Db 422 HKLGVTCTIAHALEKTYPPNSDIYLDKFDSDQYHSCQFTADLLAMNHTDPIITSTFOEI 481
Qy 486 AGSKDTGQYESHTAFTLPGLYRVVHGIDVPDKFNIVSPGADMEIYFPYTEKRLKHF 545
Db 482 AGSKDTGQYESHIAFTLPGLYRVVHGIDVPDKFNIVSPGADMSVYYPYTEDKRLTAF 541
Qy 546 HPEIEDLLYTKVENEHLCVLNDNRKPIIFTWPRLDVKNLTGLVEMCGKNPKLRELANTL 605
Db 542 HPEIEELIYSDVENSEHKEFVLKDKKPIIFSMARLDVKNMTGLVEMYGNARLRELANTL 601
Qy 606 VVVGDRRKESKDLSEKAEKMMKMFELIDKYNLNGQFRWISSQMRIRNVELYRYICDTKG 665
Db 602 VIAGDHGKESKDRREGQAEFKMYSLIDYKLGKHIRWISAQMNVRNGELYRYICDTKG 661
Qy 666 AFVQPALYEAFGLTYVEAMTGLPTFATCNGSPAEIIVHGKSGFNIDPYHGDQADILVD 725
Db 662 AFVQPAFYEAFAFGLTYIESMTGCLPTIATCHGGPAEIIYDGVSGLHIDPYHSDKADILVN 721
Qy 726 FFEKCKDPSHMDKISQGLKRIEKEYTKIYSERLTLTGVGYPWKHVSNLERESRRY 785
Db 722 FFDKCKADPSYWEISQGLQRIYKEYTKLYSERLMTLTGIVGFWKIVSNLERETRY 781
Qy 786 LEMFYALKYRKLAESVPLA 804
Db 782 IEMFYALKYRSLSAQVPLS 800

Search completed: June 2, 2004, 14:51:35
Job time : 21 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2004, 23:27:21 ; Search time 214 Seconds
(without alignments)
6807.225 Million cell updates/sec

Title: US-10-003-405-1

Perfect score: 2625
Sequence: 1 atggctgagcgtgcttcac.....tcaaaaaaaaaaaaaaaaaa 2625

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1497.8	57.1	3103	US-09-598-401C-57	Sequence 57, Appl
2	1053.4	40.1	2563	US-08-553-436A-7	Sequence 7, Appl
3	491.6	18.7	2700	US-08-684-005-1	Sequence 1, Appl
4	235.8	9.0	532	US-09-598-401C-56	Sequence 56, Appl
5	165.8	6.3	6386	US-08-483-376-1	Sequence 1, Appl
6	142	5.4	271	US-09-313-294A-1826	Sequence 1826, Ap
7	136.2	5.2	272	US-09-313-294A-1221	Sequence 1221, Ap
8	85.6	3.3	265	US-09-313-294A-1862	Sequence 1862, Ap
9	68	2.6	2236	US-09-697-367-1	Sequence 1, Appl
10	59.4	2.3	3625	US-08-356-354-3	Sequence 3, Appl
11	59.4	2.3	3625	US-08-356-354-1	Sequence 1, Appl
12	59.4	2.3	3740	US-08-778-656-1	Sequence 1, Appl
13	59.4	2.3	3740	US-08-778-656-1	Sequence 1, Appl
14	57.8	2.2	2930	US-08-356-354-5	Sequence 5, Appl
15	57.8	2.2	2930	US-08-778-656-5	Sequence 5, Appl
16	57.6	2.2	282	US-09-313-294A-427	Sequence 427, App
17	53.4	2.0	84	US-09-313-294A-7580	Sequence 7580, Ap
18	49.8	1.9	3509	US-08-175-471-6	Sequence 6, Appl
19	49.8	1.9	3509	US-09-078-862-2	Sequence 2, Appl
20	49.8	1.9	3509	US-09-866-153-12	Sequence 12, Appl
21	49.8	1.9	3509	US-09-693-467A-12	Sequence 12, Appl
22	48.2	1.8	3509	US-08-718-777-6	Sequence 6, Appl
23	48.2	1.8	3509	US-09-051-341-6	Sequence 6, Appl
24	45.4	1.7	1786	US-08-956-171E-618	Sequence 618, App
25	45.2	1.7	1445	US-09-697-367-19	Sequence 19, Appl
26	45	1.7	3539	US-08-853-948B-1	Sequence 1, Appl
27	44.8	1.7	3635	US-08-553-436A-5	Sequence 5, Appl

28	44.4	1.7	3509	2	US-08-429-054A-12	Sequence 12, Appl
29	44.2	1.7	210	3	US-09-125-984-1	Sequence 1, Appl
30	43.6	1.7	7218	1	US-08-232-463-14	Sequence 14, Appl
31	42	1.6	3524	4	US-09-697-367-15	Sequence 15, Appl
32	40.8	1.6	909	4	US-09-134-001C-334	Sequence 334, App
33	39.4	1.5	1664976	4	US-08-916-421B-1	Sequence 1, Appl
34	37.6	1.4	1158	4	US-09-543-681A-3266	Sequence 3266, Ap
35	37.4	1.4	1664976	4	US-08-916-421B-1	Sequence 1, Appl
36	37.2	1.4	1326	4	US-09-621-976-8976	Sequence 8976, Ap
37	37	1.4	1326	4	US-09-328-352-4064	Sequence 4064, Ap
38	36.8	1.4	8961	4	US-10-204-708-80	Sequence 80, Appl
39	36.6	1.4	2156	3	US-08-965-762-16	Sequence 16, Appl
40	36.6	1.4	2156	4	US-09-911-927-16	Sequence 16, Appl
41	36.6	1.4	2156	4	US-09-911-927-18	Sequence 18, Appl
42	36.6	1.4	2156	4	US-09-911-882-18	Sequence 18, Appl
43	36.6	1.4	2156	4	US-09-911-882-18	Sequence 18, Appl
44	36.6	1.4	2156	4	US-09-911-888-16	Sequence 16, Appl
45	36.6	1.4	2156	4	US-09-911-888-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-09-598-401C-57
; Sequence 57, Application US/09598401C
; Patent No. 6596925
; GENERAL INFORMATION:
; APPLICANT: Perera, J. Ranjan
; APPLICANT: Bagleton, Clare
; TITLE OF INVENTION: Rice, Stephen J.
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Expression
; FILE REFERENCE: 1100.10362
; CURRENT APPLICATION NUMBER: US/09/598, 401C
; PRIOR FILING DATE: 2000-06-20
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: PCT/NZ00/00018
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: U.S. No. 6596925 09/276,599
; PRIOR FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 3103
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-598-401C-57
Query Match 57.1%; Score 1497.8; DB 4; Length 3103;
Best Local Similarity 76.4%; Pred. No. 0;
Matches 1851; Conservative 0; Mismatches 568; Indels 3; Gaps 1;
1 ATGGCTGAGCGTGTCTCACTCGCGCTCCAGAGTCTCGAGCGTTGGATGAGACCTT 60
447 ATGGCTATGCGATGTTGATCTGCAAGCAGACGCTTGCGAGCGTTGGAGACCTTC 506
61 CTGCTCAAGAGGAGGAGATTTTGGCTTGTCTCAAGATCGAGGCGCAAGAGAGGA 120
507 TCTGCTCAAGGCGCAAGATTTGTTGCTTCTTCAAGGTTGAAGCGCAAGAGGC 566
121 ATTGTGCAACCATCTCAATTTATTTAGTTTAAGTTTAACTTCTCTGAAGAGAG 180
567 ATCTTGAGGCGCCAGATTTTGTGAGTTTGAAGCCATCTTGAAGAGAGAGCA 626
181 AAGCTCCATATGTTGATTTTGAAGTTAAGGCTAGTCAAGAGCATGTGTTG 240
627 AAGCTTTTATGAGGCGCTTTGAGAGTCTCAATTCACCTCAAGAGAGCATGTG 686
241 CTTGCATGAGGTTGACCTGTTGTTGTTCAAGAGGCTGTGTTGAGAGTACATTAG 300
687 CTTGCATGAGGTTGCTTGTGTTGTTGTTCAAGAGGCGGCGTGTGAGAGACATCG 746

QY 301 AATGTTACGCGCTTGTGTGAGAACTCACTGTGAGATCTCCACTTCAAGAA 360
DB 747 AAGTCATAGCGCTGTGTTCTTGAGCAATGAGAGTTGTGATCTGCACTTCAAGAA 806
QY 361 GAGCTTGTGAGAACTTCAATGAAACTTTGTGTAATGGAATTTGACCTTC 420
DB 807 GAGCTTGTGAGAACTTCAATGTAATTTGTGCTTGAAGCTTGAACCAATC 866
QY 421 AACTCATCATCCCCCGCCCAACTCTTCAAAATCATTTGTAATGAGTGAATCTCA 480
DB 867 ACTGCTCTTTTCCGCGCCCACTCTTCCAACTCATTTGCAATGAGCTGATCTC 926
QY 481 AATGTCACCTTTCCGCAAAATTTGTCATGACAAAGAGACATGACCTTTGCTGAA 540
DB 927 AATGCGCATCTCTCGCTAGCTCTTCCATGACAAAGAACTTGACCTCTGCTGAA 986
QY 541 TTCTCTGAGTCTCATGTCACAAAGGAGCAAACTGATGTAATGACAAATTCAGAC 600
DB 987 TTCTCTCAAGTCACTGCTACAAAGGAGCAAACTGATGTAATGCAAACTCAGAA 1046
QY 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGCAAGAGATCTTGTAACCTTACTCT 660
DB 1047 GTGTTCTCTTCCACATGCTCTGAGAAAGCGAGAGATCTGACCTCGCTCAAACTC 1106
QY 661 GAGACACCATGTGCGAAATTCGAAACCGGTTTCCAGAAATCGGTTTGAAGAGTTGG 720
DB 1107 GAGACCGGTAATCTCCAGATCTGAGCAACAGTTTCCAGGAGATCGGAGCTGAGCGGAGTGG 1166
QY 721 GGTGACACCGGAGAAAGCGGTGCTGAGATGATCCAACTCTTTGATCTTTGAGGCA 780
DB 1167 GGTGACACCGGAGAAAGCGGTGCTGAGATGATCCAACTCTTTGATCTTTGAGGCT 1226
QY 781 ACTGATCTTGCACCTTGAAGAGTTCTTGAAGAAATCCCATGATGTTCAATGTTGTG 840
DB 1227 CCGGACCGGTCATCTCTGAGAAAGTTCTTGAATGAGGTTCCCATGATCTTTCAAGTGTG 1286
QY 841 ATTCTCATCTCCCAAGGATCTTGCCTGACAAACAAATGTTTGGGATCTCCGACCGGT 900
DB 1287 ATCATGCTCTCCCAAGGATCTTGTCTCAGGACGAGCTTGTGTAATCCGATATACGGT 1346
QY 901 GGGGAGGTTGTTTACATTTGATGATCAAGTCCGAGCTTTGAGAAATGAGATCTCTCGT 960
DB 1347 GGGCAGGTTGTTTACATCTGATCAAGTCTGCGCTTGAAGAAATGTAATCTTCAAGC 1406
QY 961 ATTAAGCAAGAGATCAATCAATCACTCCGAAATCTCATTAATTAAGTCTTCTCT 1020
DB 1407 ATTAAAGCAAGAGATGATGATTTTACTCTCGATTTCTCATTAATCACTCGGCTTCTTCA 1466
QY 1021 GATGCTGTGAGCAACAACTGCGGTCAAGACTTGAAGAAATGATGAGCAAGCACTCG 1080
DB 1467 GACGCGGTGAGCAACCTGTGCGCAAGCGCTTGAAGAAATTTTGGGACGAGATCTCC 1526
QY 1081 GATATTTCTGAGATCTCTTCAAGAAAGAAATTTGTCAGAAATGATCTCAAGA 1140
DB 1527 CACATTTCTTCCGCTCTCTTCAAGAAATGAGAGGAGTCTCCGCAATGTAATTTCCGG 1586
QY 1141 TTGTAAGAAAGTCTGCGCATCTTGAAGAACTTACAGAGAGATGTTGCTCATGAATCTCC 1200
DB 1587 TTGCG---AGGTGTGCGCTTATTTGAAAGATACCTGAGATGTGCGAGGAACTTGCT 1643
QY 1201 AAAGAGTTGACGCGCAAGATCTGATCATCGGAAACAAAGGAGCGGCAATATCTG 1260
DB 1644 GAGAGTTGACGCGGAGCACTGATCTGATCATCGGAACTACAGATGAGAAACATTTGT 1703
QY 1261 GCTCTCTTGTGCACTAATTAAGGTGTCAACAGTGCACATCTGCGGCAATGCTTTGAG 1320
DB 1704 GCTCTCTTGTGCACTAATTAAGGTGTCAACAGTGTCAATTAAGGCAATGCTTCCAG 1763
QY 1321 AAGCAAAATATCAAGATTCAGATATCTAATGAGAGAGCTTGAAGCAAAATCAATTC 1380
DB 1764 AAGCAAAATATCAAGATTCAGATATCTAATGAGAGAGCTTGAAGCAAAATCAATTC 1823

QY 1381 TCTTGCCAAATTTACAGTGAATCTTTTGAATGAAACATACAGATTTTCAATGACGAT 1440
DB 1824 TCTTGCCAGTTCACTGCTGATCTTCAATGACCATGAAACACACCGGATCTTCAATTAACAGC 1883
QY 1441 ACTTTCAGAGAAATGTCAGAGAGAGAGCACTGTTGTCATTAACGAGAGCACTGCT 1500
DB 1884 ACTTTCAGAGAAATGTCAGAGAGAGAGATCACTGTTGTCATTAACGAGAGCACTGCT 1943
QY 1501 TTCACTCTTCTGATCTTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1944 TTCACTCTTCTGATCTTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 2003
QY 1561 AACTTGTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 2004 AACTTGTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2063
QY 1621 AGGTGAGAGATTTCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
DB 2064 CGGTGAGAGATTTCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2123
QY 1681 GAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 2124 GAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2183
QY 1741 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
DB 2184 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2243
QY 1801 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 2244 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2303
QY 1861 GAGAGGCTGAAATGAGAAATGTTGAGCTGATGACAACTTGAACGAGCAA 1920
DB 2304 GAGAGGCTGAAATGAGAAATGTTGAGCTGATGACAACTTGAACGAGCAA 2363
QY 1921 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
DB 2364 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2423
QY 1981 TTGAGCAAGAAAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
DB 2424 TTGAGCAAGAAAGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2483
QY 2041 GTGAGGCAATGATCTGCGGTTTGCACAACTTGCACAACTTGCACAACTTGCACAACT 2100
DB 2484 GTGAGGCAATGATCTGCGGTTTGCACAACTTGCACAACTTGCACAACTTGCACAACT 2543
QY 2101 ATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
DB 2544 ATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2603
QY 2161 GATCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
DB 2604 GATCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
QY 2221 TCCCAAGAGGCTTGAAGCAATGAGAGAGATGATGAGAGATGATGAGAGATGATGAGAG 2280
DB 2664 TCCCAAGAGGCTTGAAGCAATGAGAGAGATGATGAGAGATGATGAGAGATGATGAGAG 2723
QY 2281 CTATGACCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
DB 2724 CTATGACCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2783
QY 2341 GAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
DB 2784 GAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2843
QY 2401 GTTCCATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2422
DB 2844 GTTCCATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2865

RESULT 2

US-08-553-436A-7

Sequence 7, Application US/08553436A

Patent No. 5866790

GENERAL INFORMATION:

APPLICANT: HESSE, Holger

APPLICANT: MULLER-ROBER, Bernd

TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE

TITLE OF INVENTION: PREPARATION OF SUGAR BEET WITH CHANGED SUCROSE

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Ostrolenko, Faber, Gerb & Soffen

STREET: 1180 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: US

ZIP: 10036-8403

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/553,436A

FILING DATE: 17-NOV-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP94/01671

FILING DATE: 20-MAY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P 4317596.1

FILING DATE: 24-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: Weillman, Edward

REGISTRATION NUMBER: 24,735

REFERENCE/DOCKET NUMBER: P/951-117

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 382-0700

TELEFAX: (212) 382-0888

TELEX: 236925

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 2563 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Beta vulgaris

INDIVIDUAL ISOLATE: Saccharose synthase

IMMEDIATE SOURCE:

LIBRARY: phage lambda zap

FEATURE:

NAME/KEY: CDS

LOCATION: 3..2303

US-08-553-436A-7

Query Match

Best Local Similarity 67.8%; Pred. No. 0;

Matches 155; Conservative 0; Mismatches 717; Indels 12; Gaps 4;

164 CTGAAGAGACAGAAAGAGCTCGCTAATGTCATTTTGAAGTATGAAGCTAGTC 223
1 CTGCGAGGAGGAGAAACAAATCTTACCGATGCGCCGTTAGCGAAGTCTTAGTCTGCTC 60
224 AGAAGCAGTGTGCTGCTCATGAGGTTGCACTGCTGTTGTCAGAGCCTGAGTTT 283
61 AGAAGCAGTGTGCTGCTCATGAGGTTGCACTGCTGTTGTCAGAGCCTGAGTTT 120
284 GAGAGTACATTAAGATGATGTTCAAGCCCTTGTGTTGAGAACTCATGTTGCTGAGT 343
121 GAGAAATATGTTGTTTATATGCTCTGAACGATGAGAGCACTAATCTGCTGAGT 180

QY 1424 ATTTCATCATCAACGACTTTCAGGAATTCGAGAGACGCTGGTCAAT 1483
 DB 1258 ATTTCATCATTAACGACTTTCAGGAATTCGAGAGACGCTGGTCAAT 1317
 QY 1484 ACGAGCGACAGCTTTCAGCTTCTCTGCTCTACCGGTTGTATAGTATCGATG 1543
 DB 1318 ATGAAGCGACAGAGCTTTCAGCTTCTCTGCTCTACCGGTTGTATAGTATG 1377
 QY 1544 TGTTCGATCCCAATTCAGATTTGTTCCCTGCTGTGATATGAGATATATCTTCCCTT 1603
 DB 1378 TCTTTGATCCCAATTCAGATTTGTTCCCTGCTGTGATATGAGATATATCTTCCAT 1437
 QY 1604 ACACCGAAGAGAGCGGA--GTTGAAGCATTCATCTCGAGATCGAAGACCTTCTT 1660
 DB 1438 TTTCAGAGAGAGATGTCCTGCTCACTTCATCATGACTTATAGAGAGCGCTCCAT 1497
 QY 1661 ACACCAAGTTGAAGATGAAGAACTATGTCGTGAAGACCGCAAGCAAGCAATTC 1720
 DB 1498 TCAAACTGAGACAGAGAGAACTATGTCGTGAAGATGATGATACCTCAAGCAATTA 1557
 QY 1721 TGTTCAGATCCCAAGCTTTCAGCTTCTCTGCTGTGATATGAGATATATCTTCCCTT 1780
 DB 1558 TATTTTCATGCGAGAGCTTTCAGCTTCTCTGCTGTGATATGAGATATATCTTCCAT 1617
 QY 1781 GCAAGAACCAAGTTCGCTGATGCTTAACTTCGATGTTGATGCTG--TGATAGGC 1837
 DB 1618 GCAAGAACCAAGTTCGCTGATGCTTAACTTCGATGTTGATGCTG--TGATAGGC 1677
 QY 1838 GAAAGGATCTAAAGATTTGAGAGAGAGCTGAAATGAAGAAATGTTGAGCTGATCG 1897
 DB 1678 TAAAGGATCTAAAGATTTGAGAGAGAGCTGAAATGAAGAAATGTTGAGCTGATCG 1737
 QY 1898 ACAAGTACATTCGAG 1957
 DB 1738 AGGAGTAAATTTAAG 1797
 QY 1958 ATGTTGAACCTTTCAGATTCATTTGCGACAGAAAGTCTCTGTTGACAGCTGATGT 2017
 DB 1798 ATGTTGAACCTTTCAGATTCATTTGCGACAGAAAGTCTCTGTTGACAGCTGATGT 1857
 QY 2018 ATGAAGCTTTCAGATTCATTTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2077
 DB 1858 ATGAAGCTTTCAGATTCATTTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1917
 QY 2078 CCTGTAACGCTGAGCCAGAGAGATTCATTTGCGAGAGAGAGAGAGAGAGAGAGAG 2137
 DB 1918 CCTGTAACGCTGAGCCAGAGAGATTCATTTGCGAGAGAGAGAGAGAGAGAGAGAG 1977
 QY 2138 CTTACATGCTGATCAGCTGCTGATCATCTGCTGATTCCTTGAAGAGTGAAGAG 2197
 DB 1978 CATATCATGCTGATCA--GGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2034
 QY 2198 ATCATCTCATGCTGATCAAGATTCCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2257
 DB 2035 ATCCAACTACTGAGCTTAAATCTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGATTA 2094
 QY 2258 CATGAGAGATTTACTCGAG 2317
 DB 2095 CCTGCAAGAGATTTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2154
 QY 2318 ATGTTTCAGACCTTGAAGCGCGTGAAGAGTCTGTTACCTTGAAGATTTTATGCTT 2377
 DB 2155 ATGTTTCAGACCTTGAAGCGCGTGAAGAGTCTGTTACCTTGAAGATTTTATGCTT 2214
 QY 2378 AGTACCGTGAAGCTGCTGAATCAGTTGCATGCGAGAGAGAGATTA 2421
 DB 2215 AGTACCGTGAAGCTGCTGAATCAGTTGCATGCGAGAGAGATTA 2258

RESULT 3
 US-08-684-005-1
 ; Sequence 1, Application US/08684005
 ; Patent No. 6682918

; GENERAL INFORMATION:
 ; APPLICANT: Haselkorn, Robert
 ; APPLICANT: Buelkema, William J.
 ; APPLICANT: Bauer, Christopher C.
 ; TITLE OF INVENTION: BACTERIAL SUCROSE SYNTHASE
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: Arnold, White & Durkee
 ; STREET: P.O. Box 4433
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: United States of America
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: Concurrently Herewith
 ; APPLICATION NUMBER: US/08/684,005
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kitchell, Barbara S.
 ; REGISTRATION NUMBER: 33,928
 ; REFERENCE/DOCKET NUMBER: ARSB:508
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (512) 418-3000
 ; TELEFAX: (512) 474-7577
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2700 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-684-005-1
 ; Query Match 18.7%; Score 491.6; DB 4; Length 2700;
 ; Best Local Similarity 54.0%; Pred. No. 3.6e-146;
 ; Matches 1079; Conservative 0; Mismatches 905; Indels 15; Gaps 3;
 QY 378 TTCAATGAGAACTTTGTTGGAATTTGAGCCCTTCAACATCATTCCTCCCG 437
 DB 607 TCCTTATGAAGGCGATATATAGAACTGATTCGCGCCCTTCAACATTAACCCCAAC 666
 QY 438 CCAACCTTTCAAAATCCATGCTATGATGAGATTCCTAAATGCTACCTTGGC 497
 DB 667 CATCCGCGATCCAGAGATATGCGAAGGATGATCAATATCTCAACCTTATCTCCAG 726
 QY 498 AAATGCTTCATGACAGAGAGA--GCATGCAACCTTGTCTGATTCCTAGAGTCA 554
 DB 727 TAAACTTTTCAAGACTCGCAACATGCTGGAAGCTGTTTAAATTTCTTGGCGCTACA 786
 QY 555 TTGTCAAGAGGCAAGAACATGATGTTGAATGACAGATTCAGAACTTGAATGCTCTTCA 614
 DB 787 TAATTACAAGTGAATTCATTAATTAACATCAATTAATCAACAGCAACAATATTC 846
 QY 615 ACATGTTTGAAGAAACAGAGAGATCTTGTGATACCTTCTCGAGACACCATGTGC 674
 DB 847 ACAGCAAGTTAAABACCGCTTAACTTTGAGAGATGCGCCCAATATGAACCTTTCGA 906
 QY 675 GCAATTCGAACACCGGTTCCAGAGAAATCGTTTGAAGAGAGTTGGGTGACACCGGAGA 734
 DB 907 ACAATTCGCGCTGAATTAACAATATGAGTTTGAAGCGGGGTGAGTAAACAGCTTC 966
 QY 735 ACGGCTCTGAGATGATCAACTCTTTTGAATCTTCTGAGGCAACTGATCTTCTGAC 794
 DB 967 TCGTGTGCGGAGTACCTTAAACATTTTGAATGATTAATGATCACTTCCCAACCCCAAC 1026
 QY 795 CTTGAGAAATTCCTTGGAGAAATCCCAAGGTGTTCAATGTTGTGATCTCACTCCCA 854
 DB 1027 CTTGAGAACTTTATCTCTGCAATCCGAGATTTTCAGAATGCTCTTAAGTTTCAGCCCA 1086

QY 855 CGGATCTGCTGAGCAATGTTTGGGGTATCCCGACACCGGTGGGAGTTGTTA 914
 DB 1087 CGGTGGTTCGAGCAAGAGGGGGTTTATGTCCTCAGATCTAGTGTGTAAGTGTGTA 1146
 QY 915 CATCTGGATCAAGTCCGAGCTTTGAGAAATGAGATGC-----TCCTCGGTATAA 965
 DB 1147 GGTCTTGAACAGCTAAGATTTTGAAGAACCACTGCAAGAAAGTGCATATCTTGACGG 1206
 QY 966 GCAACAGAGACTCAACATCAACCCCTCCGATCTCTATTTATCTAGACTTTCTGATGC 1025
 DB 1207 TTTAGAGTATGAAAGCTCCAGCCCAAGTATTTATCTCAACCCGTCTGATTTCTTAATG 1266
 QY 1026 TGTGGAAACAACATGCGGTGCAAGACTTGAAGAAATGATGAGAAACAGAGACTCGGAT 1085
 DB 1267 TGACGAAACGTTTGTATCAAAAGTTAGAAAGTCTACGTAACAGAAAGCGCTGGAT 1326
 QY 1086 TCTTCGAGTACCTTCAGAACAGAAAGGAATTTGTCGAAATGATCTGAGATTTGA 1145
 DB 1327 TTTGCGGTACCTCTGGGGAGTTTAAACCAAGATGACGAAAGTGGATTTCTGAT 1386
 QY 1146 AAAAGTCTGCGATATCTTGAACCTTACACAGAGATTTGCTCATGAAATCTCCAAAG 1205
 DB 1387 CGAGTTTGGCTTATCTAGAAACCTTTCATTCATCAGAAAGAAATTTGGCAGA 1446
 QY 1206 GTTGCAGCGCACGCGATCTGATCATGGAACACAGCGGCAATATCGTCCCTC 1265
 DB 1447 ATTCAGAGTACAGCACTTAATCGTGTATTAATCTAGCGGAACTTAGTGTCTT 1506
 QY 1266 CTTCCTGCGACATTAATAGGTGTACACAGTGCACATCCGCCATCTTTGAGAGAAC 1325
 DB 1507 TCTGTGACGAGAGGAGTAAAGTTAACCAATGCAACATCGCTATAGTAAATC 1566
 QY 1326 AAAATATTCAGATTCATATCTATGGAAGAGTGAAGCAATTCATTTCTCTTG 1385
 DB 1567 CAATATCTGTTATGTAACCTCTACGCAAGATTTGAGAAATATCTCTCTT 1626
 QY 1386 CCAATTTACAGCTATCTTTTGAAGTAAACCATACATTTCAATCAACAGTACTT 1445
 DB 1627 ACAATATCGGCTATTAATAGCTATTAATGCTGTAATCTCGTCAACAGACCTTA 1686
 QY 1446 CCAGAAATTCAGAGAACAGACACTGTTGTCATATCAGAGACCACTGCTTTTAC 1505
 DB 1687 TCAAGAAATTTTGGCACACAGACAGATAGGAGAGATAGTCTTCAAAATGCTTAC 1746
 QY 1506 TCTTCTGCTCTACCGTGTGTATCATGTATGATGTGTTGATCCCAATTCACAT 1565
 DB 1747 CATGCCGGAATGATATCATGTGTCAACGCAATGAAATTTAGCCCAATTTAAGT 1806
 QY 1566 TGTTCCTGCTGTGATATGAGATATCTCCCTTACACCGAAGAGAGCGAGCTT 1625
 DB 1807 TGTACCGCTGTGTGAATGAATTTCTACTTTCTTACACCACTCAAAACAGAT 1866
 QY 1626 GAAGCATTTTCATCTGATGAGAGACCTTCTTACACCAAGTGAAGTGAAGACA 1685
 DB 1867 AGAAAGCATGCGATGCGCTTAGAGAAATGCTGTTTACCTTAGAAGTCTAGCAAT 1926
 QY 1686 CTATATGTGTCTCAATGACCGCAACAGCAATTTCTGTCACATGCGCAAGGCTTATCG 1745
 DB 1927 CTTCGCGCAATCTGACGACCAAAATTAAGCTCTTATTTCTCAATGCGCGCATTTGACG 1986
 QY 1746 TGTCAAGAACTTAAACCGAGCTGTGATGTGTGCGGCAAGAACCAAGTTGCGGAT 1805
 DB 1987 AATTAATAAATCTCAAGGTTTGGCAGAAATGCTTGTGTAAGTCAAGAAATTTGCAAGCG 2046
 QY 1806 GGTATACCTGATGTTGATGTGTGATAGCGAA---AGAAATCTAAAGATTTGGAAGA 1862
 DB 2047 TTGCACTTAATTTATGTTGACGTATAGCTGTGTATGCAAGATAGAAATAGAAAGA 2106
 QY 1863 GAAGGCTGAATGAAGAAATGTTGAGCTGATGCAAGTACCACTTGAACGCGCAAT 1922
 DB 2107 AAAAGACGAAATTCGCAAACTTTACCGGATTTATGACAAATACCACTGACATGCGCAAT 2166

QY 1923 CAGATGATATCATCTCAATGAACAGATCCGAATGTTGAACCTTACCGATACCTT 1982
 DB 2167 TCGCTGTATGATGTGTGCTTATCAAAATGATCCGCGGAAATTTATGCGCTCATTTG 2226
 QY 1983 CGACAGAAAGGTGCTTTTGTACAGCTGCTGATGTATGAAGCTTTGATGAAGTTGT 2042
 DB 2227 CGATGCGCAAGGCAATTTTGTACAGCCAGCATATTTGAAGCCCTTTGGTGAACAATCT 2286
 QY 2043 GGAGGCAATGACTGGGTTTGGCAACATTCGAACCTGTAAAGGTGAGACCGAGAGAT 2102
 DB 2287 GGAATCAATGATTTCCGATTTGCCAATGCTTACCAATTTGGGGGCCATTTGGAGAT 2346
 QY 2103 TATGTCATGGAATATGCTGTTCAACATTTGATCTTACCATGATGATCAAGCTGTGA 2162
 DB 2347 TATTCAGATTAAGTTATAGCTCTTCTTACATTTAACCTTACATCTAGAAAGAACGCAAC 2406
 QY 2163 CATCTGCTGATTTCTTTGAAAGTGAAGAAAGATTCATCTGACGGATTAAGATCTC 2222
 DB 2407 AAAATCTTGTATTTCTGACCAATGCGAAACAAATCTTAACTTATGGAACATTAATTT 2466
 QY 2223 CCAAGAGGCTTGAAGAAATGAGAGAGATATACATGAAATTTTACTCGAGAGACT 2282
 DB 2467 CGAAGAAAGCAATGACAGAGATATATGATACATACCTGAAATTAACACACATTAAGCT 2526
 QY 2283 ATGACCTGACAGAGATGATGATTTCTGAAAGCATGTTTCAACCTTGAACCGCGTGA 2342
 DB 2527 GTTAACCTTACCTGCGATTTACGCTTCTGGAATTTTAACTCGAAGAAAGCGAAGA 2586
 QY 2343 GAGTGTGTTACTTTAG 2361
 DB 2587 TTTATTAACGCTACCTTGA 2605

RESULT 4

US-09-598-401C-56
 ; Sequence 56, Application US/09598401C
 ; Patent No. 6596925
 ; GENERAL INFORMATION:
 ; APPLICANT: Perera, J. Ranjan
 ; APPLICANT: Eagleston, Clare
 ; TITLE OF INVENTION: Compositions and Methods for the
 ; FILE OF INVENTION: Modification of Gene Expression
 ; FILE REFERENCE: 11000.1036c2
 ; CURRENT APPLICATION NUMBER: US/09/598,401C
 ; PRIOR FILING DATE: 2000-06-20
 ; PRIOR APPLICATION NUMBER: U.S. No. 6596925 60/146,591
 ; PRIOR FILING DATE: 1999-07-30
 ; PRIOR APPLICATION NUMBER: PCT/N200/00018
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: U.S. No. 6596925 09/276,599
 ; PRIOR FILING DATE: 1999-03-25
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 56
 ; LENGTH: 532
 ; TYPE: DNA
 ; ORGANISM: Pinus radiata
 ; US-09-598-401C-56

Query Match 9.0%; Score 235.8; DB 4; Length 532;

Best Local Similarity 76.8%; Pred. No. 7,1e-65;

Matches 288; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 1 ATGCTGAGCGGTCTTCACTCGGCTTCAACAGTCTCCGTGAGCGTTGGATGAGACCTT 60
 DB 158 ATGCTGATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 217
 QY 61 CTGCTGACAGGAGAGAGATTTTGGCTGTCTCAAGATTCGAGGCGAAGAAAGGA 120
 DB 218 TCTGCTACCGCAACGATTTGTGCTCTTCTTCAAGGTTGAAGCGCAAGGCG 277
 QY 121 ATTCGAAACACATCAATTAATTTAGAGTTGAAGTATCTCTGAAGAGACAGAAAG 180

Db 278 ATCTTGACGCCACACAGATTCTTGTAGTTGAGGCCATCTCTGAGAGACAGAGCA 337
Qy 181 AAGCTCGGTANAGGATTTTGTGAAGTATGAAAGGCTAGTCAAGAGCATGCTGTG 240
Db 338 AAGCTTCTGATGGGCGCTTGTGTAATCTCTCAATCACTCAGAGAGCATTTGTG 397
Qy 241 CCTCATGAGGTGCACTTCTGTCTGTCGAAGCGCTGTGTTGGAGTACATTAGAGTG 300
Db 398 CCTCATGAGGTGCTCTTGTCTGTCTGTCGAAGCGCGGTGTGGAGACATCCGTGTG 457
Qy 301 AATGTCACGCCCTTGTGTGAGAACTCACTGTCTGATCTTCACATTCAGAGAA 360
Db 458 AACGTCATGCGCTGTCTGTGAGCAATTGAGAGGTGTGATCTGACATTCAGAGAA 517
Qy 361 GAGCTGTGTATGA 375
Db 518 GAGCTTGTCTATGA 532

RESULT 5

US-08-483-376-1
Sequence 1, Application US/08483376

Patent No. 595330

GENERAL INFORMATION:

APPLICANT: Vasil, Vimala

APPLICANT: Clancy, Maureen A.

APPLICANT: Ferl, Robert J.

APPLICANT: Vasil, Indra K.

APPLICANT: Hannah, L. C.

TITLE OF INVENTION: No. 595330e1 Means for Enhancing Gene

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Greenlee, Winner and Sullivan, P. C.

STREET: 5370 Manhattan Circle, Suite 201

CITY: Boulder

STATE: Colorado

COUNTRY: US

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/483,376

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/418,540

FILING DATE: 07-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/102,115

FILING DATE: 04-AUG-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/830,956

FILING DATE: 05-FEB-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/353,854

FILING DATE: 18-MAY-1989

ATTORNEY/AGENT INFORMATION:

NAME: Feider, Donna M.

REGISTRATION NUMBER: 33,878

REFERENCE/DOCKET NUMBER: 10-94B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 499-8080

TELEFAX: (303) 499-8089

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 6386 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Zea mays
STRAIN: Black Sweet
FEATURE:
NAME/KEY: exon
LOCATION: 131..182
FEATURE:
NAME/KEY: exon
LOCATION: 1211..1324
FEATURE:
NAME/KEY: exon
LOCATION: 1828..1948
FEATURE:
NAME/KEY: exon
LOCATION: 2041..2187
FEATURE:
NAME/KEY: exon
LOCATION: 2269..2460
FEATURE:
NAME/KEY: exon
LOCATION: 2605..2728
FEATURE:
NAME/KEY: exon
LOCATION: 2822..3038
FEATURE:
NAME/KEY: exon
LOCATION: 3256..3351
FEATURE:
NAME/KEY: exon
LOCATION: 3447..3620
FEATURE:
NAME/KEY: exon
LOCATION: 3702..3818
FEATURE:
NAME/KEY: exon
LOCATION: 3912..4078
FEATURE:
NAME/KEY: exon
LOCATION: 4158..4381
FEATURE:
NAME/KEY: exon
LOCATION: 4517..4835
FEATURE:
NAME/KEY: exon
LOCATION: 4768..5212
FEATURE:
NAME/KEY: exon
LOCATION: 5372..5510
FEATURE:
NAME/KEY: exon
LOCATION: 5636..5917
US-08-483-376-1

Query Match 6.3%; Score 165.8; DB 2; Length 6386;

Best Local Similarity 55.8%; Pred. No. 1.2e-41;

Matches 517; Conservative 0; Mismatches 197; Indels 213; Gaps 2;

Qy 1294 CAGTGACCAATCGCCATGCTTTGGAGAGACAAATATCCAGATTGATATCTATGG 1353

Db 3909 CAGTGACCAATCGCTATGCTTGGAGAGAAACCAATATCCCACTGACATATCTTG 3968

Qy 1354 AAGAGCTTGAAGACCAATATCCATTTCTTTGGCAATTTACAGTGTATCTTTTGCATG 1413

Db 3969 GACAAATTCACAGCCAGTACCACTTCTTGGCAGTTCAAGCTGACTTATTTGCATG 4028

Qy 1414 AACCATCAGATTTCATCATCACAGTACTTTCAGAGAAATTTGCAGAA----- 1461

Db 4029 AACCAACCAATTCATCATCACAGCAATTCAGAGAAATTCGCGGAGAGGTAGAAATTTG 4088

Db 240 AGAAGAACCGATCATCTTCTCGATGGCGCTC 272

RESULT 8
US-09-313-294A-1862

Sequence 1862, Application US/09313294A

Patent No. 6476212

GENERAL INFORMATION:

APPLICANT: Ito, Laura Y.

APPLICANT: Ito, Laura Y.

APPLICANT: Sherman, Bradley K.

TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

FILE REFERENCE: PL-0017 US

CURRENT FILING DATE: 1999-05-14

NUMBER OF SEQ ID NOS: 7600

SOFTWARE: PERL Program

SEQ ID NO 1862

LENGTH: 265

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. 6476212 700551647H1

US-09-313-294A-1862

Query Match 3.3%; Score 85.6; DB 4; Length 265;

Best Local Similarity 60.2%; Pred. No. 4.7e-17;

Matches 159; Conservative 0; Mismatches 104; Indels 1; Gaps 1;

QY 474 GTTCTTAATGTCACCTTTGGGCAAAATTTGTCATGACAGAGAGATGACACCTTT 533

Db 1 GTTCTCAACCGACACCTTCTCTCATCATGTCGCAACAGGAGATGCTTGGAGCCCTT 60

QY 534 GCTGAATTCCTCAGAGTCATTTGACAGAGGCAAGAGATGATGTAATGACAGAT 593

Db 61 GCTGATTTCTTCCTGTCGACCGGACAGAGGAGATTTATGATGCT-TATGATAGAT 119

QY 594 TCAGAACTTGAATGCTCTTCAACATGTTTGAAGAGAGAGATATCTTGATCCCT 653

Db 120 ACAAGCTTTGGGAGGCTTCAAGTGTGCTGACCAAGCTGAGGACCTTGTCAAAGCT 179

QY 654 ACCTTCGAGACACCATGTCGCAATTTGAACACCGGTTTCCAGGAATCGTTTGAAG 713

Db 180 CCTGCTGACACCATCTACTCACAATTTGCTTAAATTTCAAGAGTGGGCTGAGAA 239

QY 714 AGTTGGGGTGACACCGCAAGC 737

Db 240 GGTGGGTGATACAGAGGCGATG 263

RESULT 9
US-09-697-367-1

Sequence 1, Application US/09697367

Patent No. 6323015

GENERAL INFORMATION:

APPLICANT: Orozco Jr., Emil M.

APPLICANT: Calmi, Perry G.

APPLICANT: Meng, Zude

APPLICANT: Tarczynski, Mitchell

TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE

FILE REFERENCE: BB116 US NA

CURRENT FILING DATE: 2000-10-26

PRIOR FILING DATE: 1998-MAY-07

PRIOR APPLICATION NUMBER: PCT/US99/09865

PRIOR FILING DATE: 1999-MAY-06

NUMBER OF SEQ ID NOS: 24

SOFTWARE: Microsoft Office 97

SEQ ID NO 1

LENGTH: 2236

TYPE: DNA

ORGANISM: Catalpa speciosa
US-09-697-367-1

Query Match 2.6%; Score 68; DB 4; Length 2236;

Best Local Similarity 48.6%; Pred. No. 1e-10;

Matches 279; Conservative 0; Mismatches 285; Indels 10; Gaps 3;

QY 1612 GAGAGCGGAGGTTGAAGCATTTCCATCTGAGATGCAAGACCTTTTACACCAAGTT 1671

Db 170 GAGACATGATGATGAGAGCTGAGAGCAACAGAGTGAAGTCTCCAGAAACACTTATT 229

QY 1672 GAGAATGAGAACCTATGCTGCTCAATGACCGCAACGCAACCAATCTGTTGACATG 1721

Db 230 GGGCAGAGGTAAGCTGTTCTTTTC---AAATCAAGAGAGCTATGATTTCTGCACTT 285

QY 1732 CCAAGCTTATGCTGTCAGAAACTTAAACCGAGCTGCTGAGTGGCGGCAAGACCA 1791

Db 286 GCGAGGCGAGATCCAAAGAAAGAAACCTCACTATCTGTCAGAGATTTGGGGAAATGCGA 345

QY 1792 AAGTTGCTGATGTTGCTTAACTGCTGATGTTGATGTTGATGAGCAAGAAATCTAA 1851

Db 346 CCACTAAGGAGAGCTGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 405

QY 1852 GATTGGAAGAGAGGCT---GAAATGAGAAATGTTGAGCTGATCGACAACTACAA 1908

Db 406 ATGTGGGAACCAATGCTTCAATGCTTCAATGCTTCAATGCTTCAATGCTTCAATGCT 465

QY 1909 TTGAACGGCAATTCAGATGATATCAATCAATCAATCAATCAATCAATCAATCAATCA 1968

Db 466 CTGATGCTGATGAGGAGATATCTTAACTCAACAGCA---ACATGATGTTCTGAAAT 522

QY 1969 TACCGATATTTGCGACAGAAAGGCTCTTTGTAAGCTTGATGATGAAAGCTTT 2028

Db 523 TACCGTACAGCAAGCAAGAGGCTGTTTCAATCAATCAATCAATCAATCAATCAATCA 582

QY 2029 GATTTGACAGTTGAGAGGCAATGATGCTGCTTGGCAATTCGCAACCTGTAACGCT 2088

Db 583 GGGCTTACTCTCATGAGGCTACAGCATATGTTGCAATGTTGCAAGCAAAATGCT 642

QY 2089 GAGACAGCCAGATTAATGTCATGAGGAAATCTGTTCAACATGATCTTCAACATG 2148

Db 643 GGCCCTGTTGATATACAGAGGTTCTGAGCAATGCTCTCTGTTGATCCCAATCAG 702

QY 2149 GATCAAGCTGCTGACATGCTGCTGATTTCTTTG 2182

Db 703 CAGTCATGCTGATGCTCTTTGAACTGTTG 736

RESULT 10
US-08-356-354-3

Sequence 3, Application US/08356354

Patent No. 5767365

GENERAL INFORMATION:

APPLICANT: SONNEWALD, Uwe

TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE

TITLE OF INVENTION: PREPARATION OF PLANTS WITH CHANGED SUCROSE CONCENTRATION

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen

STREET: 1180 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: US

ZIP: 10036-8403

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/356,354

FILING DATE: 20-DEC-1994

CLASSIFICATION: 435

```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER: US PCT/EP93/01605
3      FILING DATE: 22-JUN-1993
4      PRIOR APPLICATION DATA:
5      APPLICATION NUMBER: DE P42 20 758.4
6      FILING DATE: 24-JUN-1992
7      ATTORNEY/AGENT INFORMATION:
8      NAME: Meliman, Edward A.
9      REGISTRATION NUMBER: 24,735
10     REFERENCE/DOCKET NUMBER: P/951-105
11     TELECOMMUNICATION INFORMATION:
12     TELEPHONE: (212) 382-0700
13     TELEFAX: (212) 382-0888
14     TELEX: 236925
15     INFORMATION FOR SEQ ID NO: 3:
16     SEQUENCE CHARACTERISTICS:
17     LENGTH: 3625 base pairs
18     TYPE: nucleic acid
19     STRANDEDNESS: single
20     TOPOLOGY: linear
21     MOLECULE TYPE: cDNA
22     ORIGINAL SOURCE:
23     ORGANISM: Solanum tuberosum
24     FEATURE:
25     NAME/KEY: CDS
26     LOCATION: 121..3282
27     OTHER INFORMATION: /note= "Sucrose-Phosphate-Synthase"
28
29 US-08-356-354-3

```

Query Match	2.3%	Score 59.4	DB 1	length 3655
Best Local Similarity	48.0%	Pred. No. 8.5e+08		
Matches 233	Conservative 0	Mismatches 246	Indels 6	Gaps 2

Qy	1701	TGACCGCAACAGCCAACTTCGTTCAACATGCAGACAGGCTTGATCGTCGCAAGAACTTAACTAC	1766
Db	1530	TAAATCCAGGAAGCCTATGAAATCTGCACCTTGCTAGGCTGATCCAGAAAGAACTCAC	1589
Qy	1761	CGGACTCGTCAGTGTGTCGGCCAAAGAACCCAAAGTTGGCTGAGTTGGCTTAACCTCGTAGT	1822
Db	1590	TACTTTAGTGAAGACATTTGGTGAATGTCGCCATTTTGAGAGAGCTTGCTTAATCTTACTTT	1644
Qy	1821	TGTAGTGTGTATAGCCGAAAGAACTTAAAGATTTGGAGAGAAAGGCTGA--ATGA	1877
Db	1650	GATAAATGGGTAAATGAGATTAATATGACGAATATGCTAAGACAAATTCGTGACCTTCTTCT	1709
Qy	1878	GAAATATGTTGAGCTGATGCAACAGTACACTTGAACGGCCAAATTCAGATGATATCATC	1933
Db	1710	TTCAATCTTGAATAATGATAGATTAAGTATGATCTTATAGTCAAGTAGCTTATCTTAACA	1765
Qy	1938	TCAAATGAACGAATCCGAAATGTTGAACCTTACCGATACATTTGGACACGAAAGGTGC	1997
Db	1770	CCAAACACAGTCA--GATGTTCCGTGAATATCAACGCTTTCGTGAAAGACTCAAGGGGT	1822
Qy	1998	CTTTGTACAGCTCGATGTATGAAAGCCTTTGGATTTGACAGTGTGAGAGCAATGACTTTG	2057
Db	1827	TTTTATTATATCAGCTTTTATTATGACCCTTTTGGACTGACTTTGATTTGAGGACGACGCTTA	1886
Qy	2058	CGGTTTGCCAACTTTCGCAACCTGTAAACGCTGACACAGCCGAGATTAATGTTCATGGGA	2117
Db	1887	TGGCTCTCCCAATGGTAAAGCCAAAAAATGGAGAACCTTTGATATACATPAAGGGTTCTTGA	1946
Qy	2118	ATCTGCTTTCAACATTTGATCTCTTACCATGTGTGATCAGCTGTGTCACATACCTGTCGATTT	2177
Db	1947	CAATGGTCTCTTAATGTGATCCCAATGATCAGCAGCAATTCGTGATGCTCTTTGAAAGTT	2006
Qy	2178	CTTTG 2182	
Db	2007	GGTTG 2011	

RESULT 11
US-08-778-656-3
; Sequence 3, Application US/08778656

Patent No. 5976869
GENERAL INFORMATION:
APPLICANT: SONNEMALD, Uwe
TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE
TITLE OF INVENTION: PREPARATION OF PLANTS WITH CHANGED SUCROSE CONCENTRATION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/778,656
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/356,354
FILING DATE: 20-DEC-1994
APPLICATION NUMBER: US PCT/EP93/01605
FILING DATE: 22-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P42 20 758.4
FILING DATE: 24-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meliman, Edward A.
REGISTRATION NUMBER: 24,735
REFERENCE/DOCKET NUMBER: P/951-105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 382-0700
TELEFAX: (212) 382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3625 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Solanum tuberosum
FEATURE:
NAME/KEY: CDS
LOCATION: 121..3282
OTHER INFORMATION: /note= "Sucrose-Phosphate-Synthase"
US-08-778-656-3

Query Match	2.3%	Score 59.4;	DB 2;	Length 3625;
Best Local Similarity	48.0%;	Pred. No. 8.5e-08;		
Matches 233; Conservative	0;	Mismatches 246;	Indels 6;	Gaps 27

QY	1701	TGACCGCAACAGCGCAATTCTGTTCACATATGCCAAGCTTGATGTGTCAAGAACTTAACT	1766
Db	1530	TAAATCCAAAGGAGGCTATGATATCTCGCATCTTGATAGGCTGATATCCAAAGAAAGCTCAC	1589
QY	1761	CGAGCTGTGAGTGGTGGCGGCAAGAACCCAAAGTTGGTGGATGTGGTAACTCTGAGT	1822
Db	1590	TACTTATAGTGAAAGCAATTTGTGTGATGTCTGTCACATTGAGAGAGCTTGTCTAATCTTACTTT	1644
QY	1821	TGTAGTGTGTGATAGCGAAAGCAATCTAAAGATTTGGAAGAGAAGCTGAA--ATGAA	1877
Db	1650	GATATATGGGTATATCGATATATATATCGAAGAAATGTCTAGCACCAATCTCGACTTCTTCT	1709
QY	1878	GAAATATGTTGAGCTGATCGACAGATACACTTGAAAGGCGCAATTCCAGATGAGATATATCT	1937
Db	1710	TTCTATCTTTGAAATATGATATAGTATGATCTTTATAGTGTCAAGTACTTATCTTAACA	1768

QY 1938 TCAATGACAGATCCGAATGTTGAACCTTACCGATACATTTGCCACAGAAAGTGC 1997
DB 1770 CCACAGAGATCA---GATGTTCTGATATCTACCGTCTGCTGCAAGACTTAAGGGTGT 1826
QY 1998 CTTTGATCAGCTGATGATGATGAGCCTTTGATTTGACAGTTGAGGAGCAATGACTTG 2057
DB 1827 TTTTATTAATCAGCTTTTATTTAGACCTTTTGAGCTGACTTGTATGATGAGGAGCACTTA 1886
QY 2058 CGGTTTGCACATTCGCAACCTGTAACGCTGAGACAGCCGAGATTATTTGCAATGGAA 2117
DB 1887 TGGTCTCCCAATGATGACCAAAAAATGAGAGACCTGTGATATATACATAGGGTCTTGA 1946
QY 2118 ATGCGTTTCAACATGATGCTTACCATGATGATACAGCTGCTGACATCTGCTGATTT 2177
DB 1947 CAATGCTCTTATGATGATCCCATGATCAGCAGGCAATGCTGATCTTTTGAAGTT 2006
QY 2178 CTTTG 2182
DB 2007 GGTGG 2011

RESULT 12

US-08-356-354-1
; Sequence 1, Application US/08356354
; Patent No. 5767365

GENERAL INFORMATION:

APPLICANT: SONNEMALD, Uwe

TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ostroienk, Faber, Gerb & Soffen

STREET: 1180 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: US

ZIP: 10036-8403

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/356,354

FILING DATE: 20-DEC-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US PCT/EP93/01605

FILING DATE: 22-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P42 20 758.4

FILING DATE: 24-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Weilmann, Edward A.

REGISTRATION NUMBER: 24,735

REFERENCE/DOCKET NUMBER: F/951-105

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 382-0700

TELEFAX: (212) 382-0888

TELEX: 236925

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3740 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Solanum tuberosum

FEATURE:

NAME/KEY: CDS

LOCATION: 957..3494

OTHER INFORMATION: /note= "Sucrose-phosphate-Synthase"

US-08-356-354-1

Query Match 2.3%; Score 59.4; DB 1; Length 3740;

Best Local Similarity 48.0%; Pred. No. 8.7e-08;

Matches 233; Conservative 0; Mismatches 246; Indels 6; Gaps 2;

QY 1701 TGACCGAAGAACGCAATTCGTTCACATGCAAGCTTGATGCTGATCAAGAACTTAAC 1760
DB 1742 TAATCCAGAGAGCTATGATGATCTGCACTTGCTAGGCTGATCCAGAGAACTCAC 1801
QY 1761 CGGACTGCTGAGTGTGCGGCAAGAACCCAAAGTTGCTGATGCTGCTAATCTGATG 1820
DB 1802 TACTTTAGTGAAGCAATTTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1861
QY 1821 TGTAGTGTGATGAGCGGAAGAAATCTAAAGATTGGAAGAGAGCTGAA--ATGAA 1877
DB 1862 GATATAGGTATATCGATATATATCGAAGAAATGCTAGCAACCAATTCGCACTTCTCT 1921
QY 1878 GAAATGTTGAGCTGATCGCAAGTACAACTTGAAAGCGCAATTCAGATGATATCATC 1937
DB 1922 TTCAATCTTGAATAATGATATGATGATGATGATGATGATGATGATGATGATGAT 1981
QY 1938 TCAATGACAGATCCGAATGTTGAACCTTACCGATACATTTGCGACAGAAAGTGC 1997
DB 1982 CCACAGAGATCA---GATGTTCTGATATCTACCGTCTGCTGCAAGACTTAAGGGTGT 2038
QY 1998 CTTTGATCAGCTGATGATGATGAGCCTTTGATTTGACAGTTGAGGAGCAATGACTTG 2057
DB 2039 TTTTATTAATCAGCTTTTATTTAGACCTTTTGAGCTGACTTGTATGATGAGGAGCACTTA 2098
QY 2058 CGGTTTGCACATTCGCAACCTGTAACGCTGAGACCGGAGATTATTTGCCATGGAA 2117
DB 2099 TGGTCTCCCAATGATGACCAAAAAATGAGAGACCTGTGATATATCATGAGGTTCTTGA 2158
QY 2118 ATGCGTTTCAACATGATGCTTACCATGATGATACAGCTGCTGACATCTGCTGATTT 2177
DB 2159 CAATGCTCTTATGATGATCCCATGATCAGCAGGCAATGCTGATGCTCTTTGAAGTT 2218
QY 2178 CTTTG 2182
DB 2219 GGTGG 2223

RESULT 13

US-08-778-656-1
; Sequence 1, Application US/08778656
; Patent No. 5976869

GENERAL INFORMATION:

APPLICANT: SONNEMALD, Uwe

TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ostroienk, Faber, Gerb & Soffen

STREET: 1180 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: US

ZIP: 10036-8403

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/778,656

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/356,354

FILING DATE: 20-DEC-1994

APPLICATION NUMBER: US PCT/EP93/01605

FILING DATE: 22-JUN-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P42 20 758.4
FILING DATE: 24-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meliman, Edward A.
REGISTRATION NUMBER: 24,735
REFERENCE/DOCKET NUMBER: P/951-105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 382-0700
TELEFAX: (212) 382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3740 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORGANISM: Solanum tuberosum
FEATURE:
NAME/KEY: CDS
LOCATION: 957..3494
OTHER INFORMATION: /note= "Sucrose-Phosphate-Synthase"
US-08-778-656-1

Query Match 2.3%; Score 59.4; DB 2; Length 3740;
Best Local Similarity 48.0%; Pred. No. 8.7e-08;
Matches 233; Conservative 0; Mismatches 246; Indels 6; Gaps 2;
QY 1701 TGACCGCAACAGCAATTCCTGTTCAATGCAAGCTTGATGTCAGAACTTAC 1760
DB 1742 TAATCGAAGAGAGCTATGATGCTGCACTTGATGCTGATCCAGAAAGACCTAC 1801
QY 1761 CGAAGCTGTCAGTGTGCGCGCAAGAACCCAAAGTTCGTGATGCTGTAAGT 1820
DB 1802 TACTTGTAGTAAGCATTTGGTGAATGTCGTCATGAGAGAGCTTGCTTAATCTT 1861
QY 1821 TGTAGGTGCTATAGCGCAAGATCTTAAGATTGGAAGAGAGCTGAA--ATGAA 1877
DB 1862 GATATAGGTATGAGATTAATATGACGAAGATGCTAGCAACATTTGCACTTCTT 1921
QY 1878 GAAATGTTGAGCTGATGCAAGATCAACTTGAACGGCAATTCAGATGATATCATC 1937
DB 1922 TTCATCTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATG 1981
QY 1938 TCAATGAAAGAAATCCGAAATGTTGAACCTTACCGATACATTTGCGACAGAAAGTGC 1997
DB 1982 CCACAGAGAGTCA--GATGTTCTGATATCTACCGTCTGCGAAGAGCTAAGGGTGT 2038
QY 1998 CTTTGTACAGCTGATGATGAAAGCTTTGATGACAGTTGATGAGGCAATGACTTG 2057
DB 2039 TTTTATTAATCAGCTTTTATTTAGCCCTTTTGAAGCTGATGATGAGGAGCACTTA 2098
QY 2058 CGGTTTGCAATTCGCAACCTGTAACGATGAGAGCGAGCAAGATTAATGTCATGGGAA 2117
DB 2099 TGGTCTCCCAATGAGTACCAAAAAATGAGAGACCTTTGATATACATAGGGTTCTTGA 2158
QY 2118 ATCTGTTTCAACATGATGATCTTACATGATGATGATGATGATGATGATGATG 2177
DB 2159 CAATGATCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 2218
QY 2178 CTTTG 2182
DB 2219 GGTG 2223

RESULT 14
US-08-356-354-5
Sequence 5, Application US/08356354
Patent No. 5767365
GENERAL INFORMATION:
APPLICANT: SONNEWALD, Uwe

TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE
PREPARATION OF PLANTS WITH CHANGED SUCROSE CONCENTRATION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostroent, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,354
FILING DATE: 20-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/EP93/01605
FILING DATE: 22-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P42 20 758.4
FILING DATE: 24-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meliman, Edward A.
REGISTRATION NUMBER: 24,735
REFERENCE/DOCKET NUMBER: P/951-105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 382-0700
TELEFAX: (212) 382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2930 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Solanum tuberosum
FEATURE:
NAME/KEY: CDS
LOCATION: 118..2841
OTHER INFORMATION: /note= "Sucrose-Phosphate-Synthase"
US-08-356-354-5

Query Match 2.2%; Score 57.8; DB 1; Length 2930;
Best Local Similarity 47.8%; Pred. No. 2.3e-07;
Matches 232; Conservative 0; Mismatches 247; Indels 6; Gaps 2;
QY 1701 TGACCGCAACAGCAATTCCTGTTCAATGCAAGCTTGATGTCAGAACTTAC 1760
DB 1089 TAATCGAAGAGAGCTATGATGCTGCACTTGATGCTGATCCAGAAAGACCTAC 1148
QY 1761 CGAAGCTGTCAGTGTGCGCGCAAGAACCCAAAGTTCGTGATGCTGTAAGT 1820
DB 1149 TACTTGTAGTAAGCATTTGGTGAATGTCGTCATGAGAGACCTTCTAATCTTACTT 1208
QY 1821 TGTAGGTGATAGCGGAAGAAATCTAAAGATTGGAAGAGAGCTGAA--ATGAA 1877
DB 1209 GATATAGGTATGAGATTAATATGACGAAGATGCTAGCAACATTTGCACTTCTTCT 1268
QY 1878 GAAATGTTGAGCTGATGCAAGATCAACTTGAACGGCAATTCAGATGATATCATC 1937
DB 1269 TTCATCTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATG 1328
QY 1938 TCAATGAAAGAAATCCGAAATGTTGAACCTTACCGATACATTTGCGACAGAAAGTGC 1997
DB 1329 CCACAGAGAGTCA--GATGTTCTGATATCTACCGTCTGCGAAGAGCTAAGGGTGT 1385
QY 1998 CTTTGTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2057

Db 1386 TTTATTATATCAGCTTTTATTATGACCTTTTGGACTGCTTTGATGAGGACGACTTA 1445
Qy 2058 CGGTTTGCCAACTTCGCACTGTGAACGCTGACCAAGCATTTATTCATGCGAA 2117
Db 1446 TGGCTCCCAATGATGAGCAAAAATGAGAGACCTTTGATATACATAGGTTCTTGA 1505
Qy 2118 ATCTGTTTCAACATTTATCTTACATGATGATCAAGCTCTGACATCTGCTGATT 2177
Db 1506 CAATGCTCTTATGATGATCCCATGATCAGCAGGCAATGCTGATGCTTTTGAAGTT 1565
Qy 2178 CTTTG 2182
Db 1566 GGTGG 1570

RESULT 15
US-08-778-656-5

Sequence 5, Application US/08778656
Patent No. 5976869

GENERAL INFORMATION:

APPLICANT: SONNEMALD, Uwe

TITLE OF INVENTION: DNA SEQUENCES AND PLASMIDS FOR THE

NUMBER OF SEQUENCES: 6 PREPARATION OF PLANTS WITH CHANGED SUCROSE CONCENTRATION

ADDRESS: Ostroienk, Faber, Gerb & Soffen

STREET: 1180 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: US

ZIP: 10036-8403

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/778.656

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/356,354

FILING DATE: 20-DEC-1994

APPLICATION NUMBER: US PCT/EP93/01605

FILING DATE: 22-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P42 20 758.4

FILING DATE: 24-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: Meilman, Edward A.

REGISTRATION NUMBER: 24,735

REFERENCE/DOCKET NUMBER: P/951-105

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 382-0700

TELEFAX: (212) 382-0888

TELEX: 236925

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 2930 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Solanum tuberosum

FEATURE:

NAME/KEY: CDS

LOCATION: 118..2841

OTHER INFORMATION: /note= "Sucrose-Phosphatase-Synthase"

US-08-778-656-5

Query Match 2.2%, Score 57.8, DB 2, Length 2930,

Best Local Similarity 47.8%; Pred. No. 2,3e-07;
Matches 232; Conservative 0; Mismatches 247; Indels 6; Gaps 2;
Qy 1701 TGACCGCAACAGCAATTTGTTACAAATGCCAGGCTTGATGTCAGAACTTAAC 1760
Db 1089 TAATCCAAAGAGGCTTATGATACCTGCACTTGAGCTGATCCAAAGAACTCTAC 1148
Qy 1761 CGGACTGTGAGTGGTGGCGCAAGAACCAAAAGTTCCGTGAGTGGCTAATCTGAGT 1820
Db 1149 TACTTTGTAAGCAATTTGATGATGTCATTCATGAGAGACCTTCTAATCTTACTT 1208
Qy 1821 TGTAGTGTGATAGCGGAAGGAATCTAAAGATTGGAAGAGAGGCTGAA--ATGA 1877
Db 1209 GATATGGGTAATGAGATATATATCGAGAAATGCTAGACCAATTCGACTTCTTCT 1268
Qy 1878 GAAATGTTGAGCTGATCGACAAATGCACTTGAACGCGCAATTCAGATGATATCATC 1937
Db 1269 TTCAATCTTGAAGATGATATGATATGATATGATATGATATGATATGATATGATAT 1328
Qy 1938 TCAATGACGAATCCGAATGTTGATGATGATGATGATGATGATGATGATGATGATG 1997
Db 1329 CCAACAGCAGTCA--GATGTTCTGATATCTACCGTCTTGCCAAAGACTTAAGGGTGT 1385
Qy 1998 CTTTGTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2057
Db 1386 TTTATTATATCAGCTTTTATTATGAGCTTTTGGACTGACTTGAATGAGGACGACTTA 1445
Qy 2058 CGGTTTGCCAACTTCGCACTGTGAACGCTGACCAAGCATTTATTCATGCGAA 2117
Db 1446 TGGCTCCCAATGATGAGCAAAAATGAGAGACCTTTGATATACATAGGTTCTTGA 1505
Qy 2118 ATCTGTTTCAACATTTATCTTACATGATGATCAAGCTCTGACATCTGCTGATT 2177
Db 1506 CAATGCTCTTATGATGATCCCATGATCAGCAGGCAATGCTGATGCTTTTGAAGTT 1565
Qy 2178 CTTTG 2182
Db 1566 GGTGG 1570

Search completed: June 5, 2004, 08:36:14
Job time : 223 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 5, 2004, 00:49:26 ; Search time 1535 Seconds
(without alignments)
7601.459 Million cell updates/sec

Title: US-10-003-405-1

Perfect score: 2625
Sequence: 1 atggctgagcgtgtcttcac.....tcaaaaaaaaaaaaaaa 2625

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 295936 seqs, 2280998010 residues

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2624	100.0	2625	14	US-10-003-405-1 Sequence 1, Appl1
2	1602.6	60.1	5069	13	US-10-424-599-115562 Sequence 115562, Ap
3	1591.6	60.6	2601	13	US-10-425-114-9604 Sequence 9604, Ap
4	1590	60.6	3101	13	US-10-424-599-12144 Sequence 12144, A
5	1570.4	59.8	2772	13	US-10-425-114-6460 Sequence 6460, Ap
6	1570.4	59.8	3360	13	US-10-424-599-11458 Sequence 11458, A
7	1568	59.7	3438	13	US-10-424-599-11459 Sequence 11459, A
8	1566.4	59.7	2772	13	US-10-425-114-9594 Sequence 9594, Ap
9	1499.4	57.1	2906	16	US-10-393-840-139 Sequence 139, Appl
10	1497.8	57.1	2913	16	US-10-393-840-15 Sequence 15, Appl
11	1497.8	57.1	3103	15	US-10-137-036-57 Sequence 57, Appl
12	1486.6	56.6	2427	11	US-09-938-842A-1620 Sequence 1620, Ap
13	1486.6	56.6	2427	11	US-09-938-842A-1620 Sequence 1620, Ap
14	1331	50.7	2279	13	US-10-425-114-9355 Sequence 9355, Ap

15	1305.4	49.7	2714	15	US-10-289-757-133 Sequence 133, App
16	1305.4	49.7	2716	15	US-10-289-757-12 Sequence 12, Appl
17	1296.6	49.4	2825	15	US-10-289-757-15 Sequence 15, Appl
18	1296.4	49.3	2754	15	US-10-289-757-13 Sequence 13, Appl
19	1287.6	49.1	2957	15	US-10-289-757-134 Sequence 134, Appl
20	1281	48.8	2908	15	US-10-080-114A-6 Sequence 6, Appl1
21	1277.8	48.7	2986	13	US-10-425-114-2283 Sequence 2283, Ap
22	1277.8	48.7	2986	13	US-10-425-114-26430 Sequence 26430, A
23	1277.8	48.7	2992	13	US-10-425-114-5310 Sequence 5310, Ap
24	1277.8	48.7	3419	15	US-10-425-114-17596 Sequence 17596, A
25	1274.8	48.6	2950	15	US-10-289-757-14 Sequence 14, Appl
26	1260.4	48.0	2671	13	US-10-425-114-31764 Sequence 31764, A
27	1260.4	48.0	2817	13	US-10-425-114-33573 Sequence 33573, A
28	1254	47.8	2746	15	US-10-080-114A-4 Sequence 4, Appl1
29	1233.2	47.0	2127	13	US-10-425-114-10123 Sequence 10123, A
30	1219	46.4	2638	13	US-10-425-114-2574 Sequence 2574, Ap
31	1202.4	45.8	5877	16	US-10-260-238-25 Sequence 25, Appl
32	1190	45.3	2038	13	US-10-425-114-6461 Sequence 6461, Ap
33	1172	44.6	2832	13	US-10-425-114-15113 Sequence 15113, A
34	1150	43.8	3191	13	US-10-424-599-98889 Sequence 98889, A
35	1142.4	43.5	2505	13	US-10-425-114-35609 Sequence 35609, A
36	1115	42.5	2430	9	US-09-938-842A-2415 Sequence 2415, Ap
37	1115	42.5	2430	11	US-09-938-842A-2415 Sequence 2415, Ap
38	1104.8	42.1	2394	15	US-10-217-939-29 Sequence 29, Appl
39	1104.6	42.1	2757	15	US-10-080-114A-11 Sequence 11, Appl
40	1097	41.8	2418	15	US-10-217-939-27 Sequence 27, Appl
41	1096.2	41.8	2737	15	US-10-080-114A-1 Sequence 1, Appl1
42	1041.8	39.7	2210	13	US-10-425-114-8902 Sequence 8902, Ap
43	1031.6	39.3	2307	13	US-10-425-114-35558 Sequence 35558, A
44	982.8	37.4	2035	13	US-10-425-114-5608 Sequence 5608, Ap
45	923.6	35.2	2275	13	US-10-425-114-6036 Sequence 6036, Ap

ALIGNMENTS

RESULT 1
US-10-003-405-1
Sequence 1, Application US/10003405
Publication No. US20020116736A1
GENERAL INFORMATION:
APPLICANT: Ruan, Yong-Ling
APPLICANT: Furbank, Robert T.
APPLICANT: Danny, Llewellyn J.
TITLE OF INVENTION: Modification of sucrose synthase gene expression in plant tissue
FILE REFERENCE: GHSUSY WO1
CURRENT FILING DATE: 2000-12-18
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: 60/251852
PRIOR FILING DATE: 2000-12-08
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 2625
TYPE: DNA
ORGANISM: Gossypium hirsutum
FEATURE:
NAME/KEY: misc feature
LOCATION: (1240)..
OTHER INFORMATION: n = any nucleotide (a,g,c,t)
NAME/KEY: CDS
LOCATION: (1)..
OTHER INFORMATION:
US-10-003-405-1
Query Match 100.0%; Score 2624; DB 14; Length 2625;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 ATGGCTGAGCGTGTCTTCACCTCGGCTCACAGTCTCCGTTGGATGAGACCTT 60
1 ATGGCTGAGCGTGTCTTCACCTCGGCTCACAGTCTCCGTTGGATGAGACCTT 60

QY 61 CTGGCTCACAGGACGAGATTTTGGCTTGCTCTCAAGGATCGAGGGCAAGAAAAGCA 120
DB 61 CTGGCTCACAGGACGAGATTTTGGCTTGCTCTCAAGGATCGAGGGCAAGAAAAGCA 120
QY 121 ATTTCGCAACCCATCAAAATTAATCTAGAGTTTGAAGCTATCCCGAAGAGAACGAAG 180
DB 121 ATTTCGCAACCCATCAAAATTAATCTAGAGTTTGAAGCTATCCCGAAGAGAACGAAG 180
QY 181 AAGCTCGCTAATGTCATATTTTGAAGTGAAGAGCTAGTCAGAAAGCGATCGTGTG 240
DB 181 AAGCTCGCTAATGTCATATTTTGAAGTGAAGAGCTAGTCAGAAAGCGATCGTGTG 240
QY 241 CCTCATGAGGTGCACTTGCTGTCTCAAGGCTGTGTGTTGGAGTACATTAGAGTG 300
DB 241 CCTCATGAGGTGCACTTGCTGTCTCAAGGCTGTGTGTTGGAGTACATTAGAGTG 300
QY 301 AATGTCACGCGCTGTGTTGAGAGACCTCATGTGTGATCTCACTTCAAGAA 360
DB 301 AATGTCACGCGCTGTGTTGAGAGACCTCATGTGTGATCTCACTTCAAGAA 360
QY 361 GAGCTTGTGATGAAAGTTCAAAATGAAAATTGTTTGAATGGAATTTGAAGCCCTTC 420
DB 361 GAGCTTGTGATGAAAGTTCAAAATGAAAATTGTTTGAATGGAATTTGAAGCCCTTC 420
QY 421 AACTCATATTCCTCCGCGCCAACTCTTCAAAATCATTGTGTAATGTGTGAGTTCCTA 480
DB 421 AACTCATATTCCTCCGCGCCAACTCTTCAAAATCATTGTGTAATGTGTGAGTTCCTA 480
QY 481 AATGTCACCTTTGGGCAAAATTTGTCATGACAGAGAGATCCACCTTTGTCGAA 540
DB 481 AATGTCACCTTTGGGCAAAATTTGTCATGACAGAGAGATCCACCTTTGTCGAA 540
QY 541 TTCCCTCAGAGTCATTTGTCACAAAGGCAAGAACATGATGTAATGACAGAAATTCAGAAC 600
DB 541 TTCCCTCAGAGTCATTTGTCACAAAGGCAAGAACATGATGTAATGACAGAAATTCAGAAC 600
QY 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGAGAGATCTTGTATACCTTACTCT 660
DB 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGAGAGATCTTGTATACCTTACTCT 660
QY 661 GAGACACCATGTGCGGAATTCGAACAACGGGTTCGAGAAATCGGTTTGAAGAGGTGG 720
DB 661 GAGACACCATGTGCGGAATTCGAACAACGGGTTCGAGAAATCGGTTTGAAGAGGTGG 720
QY 721 GATGACACCGAGAAACGGGTGCTGAGATGATCCAACTCTTTTGAATCTTCTTGAAGCA 780
DB 721 GATGACACCGAGAAACGGGTGCTGAGATGATCCAACTCTTTTGAATCTTCTTGAAGCA 780
QY 781 ACTGATCCTTGCACCCCTTGAGAAAGTTCTTGGAGAAATCCCATGAGTTCATGTTGTG 840
DB 781 ACTGATCCTTGCACCCCTTGAGAAAGTTCTTGGAGAAATCCCATGAGTTCATGTTGTG 840
QY 841 ATTCTCATCTCCCGACGAGATCTTGCTCAAGACAAATGTTTGGGATATCCCGACACGGGT 900
DB 841 ATTCTCATCTCCCGACGAGATCTTGCTCAAGACAAATGTTTGGGATATCCCGACACGGGT 900
QY 901 GGGCAGGTGTTTACATCTTGGATCAAGTCCGAGGTTTGAAGATGAGATGCTCTCCGT 960
DB 901 GGGCAGGTGTTTACATCTTGGATCAAGTCCGAGGTTTGAAGATGAGATGCTCTCCGT 960
QY 961 ATTAAGCAACAAGACTCAACATCACCCCTCGAATCTCATTTATTAAGACTTCTTCT 1020
DB 961 ATTAAGCAACAAGACTCAACATCACCCCTCGAATCTCATTTATTAAGACTTCTTCT 1020
QY 1021 GATGCTGTCGGAACAACATGCGGTCAAGACTTGAAGAAATGAGATGAGATGCTCTCCGT 1080
DB 1021 GATGCTGTCGGAACAACATGCGGTCAAGACTTGAAGAAATGAGATGAGATGCTCTCCGT 1080
QY 1081 GATATTTCTGAGTACCTTTCAGAACAGAAAAGGAAATGTTTCAAAATGATCTCAGAA 1140
DB 1081 GATATTTCTGAGTACCTTTCAGAACAGAAAAGGAAATGTTTCAAAATGATCTCAGAA 1140

QY 1141 TTTGAAAAGCTCTGCCCATTACTTGAACCTTACACAGAGATGTTGCTCATGAATCTCC 1200
DB 1141 TTTGAAAAGCTCTGCCCATTACTTGAACCTTACACAGAGATGTTGCTCATGAATCTCC 1200
QY 1201 AAAGAGTTGACGGCAGCGCATCTGATCATCGGAAACNACAGGACGGCAATTCGTC 1260
DB 1201 AAAGAGTTGACGGCAGCGCATCTGATCATCGGAAACNACAGGACGGCAATTCGTC 1260
QY 1261 GCCTCTTGTCTGACATTAATTAAGTGTCAACAGTGCACCATCGCCATGCTTTGGAG 1320
DB 1261 GCCTCTTGTCTGACATTAATTAAGTGTCAACAGTGCACCATCGCCATGCTTTGGAG 1320
QY 1321 AAGCAAAATATTCAGATTCAGATATCTATTGAAAGAACTTGAAGCAAAATACCATTC 1380
DB 1321 AAGCAAAATATTCAGATTCAGATATCTATTGAAAGAACTTGAAGCAAAATACCATTC 1380
QY 1381 TCTTGCCAAATTAAGGCTATCTTTTGCAATGAAACCATACAGTTTCAATACCAAGT 1440
DB 1381 TCTTGCCAAATTAAGGCTATCTTTTGCAATGAAACCATACAGTTTCAATACCAAGT 1440
QY 1441 ACTTTCAGAAATTTGACAGAGAGACACTGTGTGTCATATGAGAGCCACACTGCT 1500
DB 1441 ACTTTCAGAAATTTGACAGAGAGACACTGTGTGTCATATGAGAGCCACACTGCT 1500
QY 1501 TTCACTCTTCTGATCTCTACCGTGTGTATCATGATGATGATGATGATGATGATGATG 1560
DB 1501 TTCACTCTTCTGATCTCTACCGTGTGTATCATGATGATGATGATGATGATGATGATG 1560
QY 1561 AACATGTTTCCCTGAGTGTATGAGATATATCTTCCCTTACCGGAAAGAAAGCGG 1620
DB 1561 AACATGTTTCCCTGAGTGTATGAGATATATCTTCCCTTACCGGAAAGAAAGCGG 1620
QY 1621 AGGTTGAAGCATTTCCATCTGAGATGAGAGACCTTCTTACACAAAGTTGAATGAA 1680
DB 1621 AGGTTGAAGCATTTCCATCTGAGATGAGAGACCTTCTTACACAAAGTTGAATGAA 1680
QY 1681 GAAACATTAATGTGTCTCATATGACCGCAACAGCCAAATCTGTTTCAATGCCAAGCTT 1740
DB 1681 GAAACATTAATGTGTCTCATATGACCGCAACAGCCAAATCTGTTTCAATGCCAAGCTT 1740
QY 1741 GATGTCGCAAGAACTTAAACCGGACTGTGAGTGTGCGGCAAGAAACCAAAGTGTG 1800
DB 1741 GATGTCGCAAGAACTTAAACCGGACTGTGAGTGTGCGGCAAGAAACCAAAGTGTG 1800
QY 1801 GAGTTGGCTAACTCGTATGTTGATGATGAGGAGGAAAGAAATCTTGAATTTGGAA 1860
DB 1801 GAGTTGGCTAACTCGTATGTTGATGATGAGGAGGAAAGAAATCTTGAATTTGGAA 1860
QY 1861 GAGAGGCTGAATGAAGAAATGTTGAGCTGATCAAGATCAACTTGAACGGCCAA 1920
DB 1861 GAGAGGCTGAATGAAGAAATGTTGAGCTGATCAAGATCAACTTGAACGGCCAA 1920
QY 1921 TTCAATGATGATATCTCAAAATGAACGAATCCGAATGTTGAACCTTACCGATACCTT 1980
DB 1921 TTCAATGATGATATCTCAAAATGAACGAATCCGAATGTTGAACCTTACCGATACCTT 1980
QY 1981 TGCGACAGAAAGGTGCTTGTATACAGCTGATGATGAGAGCTTGTGATGACAGTT 2040
DB 1981 TGCGACAGAAAGGTGCTTGTATACAGCTGATGATGAGAGCTTGTGATGACAGTT 2040
QY 2041 GTGAGGCAATGATCTGCGGTTTGCCAACTTTCGCAACTGTATACGGTGGACCGGAG 2100
DB 2041 GTGAGGCAATGATCTGCGGTTTGCCAACTTTCGCAACTGTATACGGTGGACCGGAG 2100
QY 2101 ATTAATGTCATGGAATCTGTTTCAACATTTGATCTTCAATGATGATGATGATGATG 2160
DB 2101 ATTAATGTCATGGAATCTGTTTCAACATTTGATCTTCAATGATGATGATGATGATG 2160
QY 2161 GACATATCTGTCGATTTCTTTGAAAAGTGAAGAAATCATCTCATCGGATTAAGATC 2220
DB 2161 GACATATCTGTCGATTTCTTTGAAAAGTGAAGAAATCATCTCATCGGATTAAGATC 2220
QY 2221 TCCCAAGAGGCTTGAACGAATGAGAGAGATATCATGAGAAATTTACTCGAAGAGA 2280

Db 2221 TCCGAGGAGGCTTGAACGATAGAGAGAGATATACATGGAATTTACTCGGAGAGA 2280
Qy 2281 CTATTGACCCCTGACAGAGAGTATAGATTTCTGGAAGCATGTTTCCAACTTGAAGCCGT 2340
Db 2281 CTATTGACCCCTGACAGAGAGTATAGATTTCTGGAAGCATGTTTCCAACTTGAAGCCGT 2340
Qy 2341 GAGAGTGTCTGTTACCTGAGATGTTTATGCTTTAAGTACCGTACCTGAGCTGATCA 2400
Db 2341 GAGAGTGTCTGTTACCTGAGATGTTTATGCTTTAAGTACCGTACCTGAGCTGATCA 2400
Qy 2401 GTTCATTTGGACAGAGAGATTAATGACCTGTTAAATACATTTGGCCGCTTTTCTTG 2460
Db 2401 GTTCATTTGGACAGAGAGATTAATGACCTGTTAAATACATTTGGCCGCTTTTCTTG 2460
Qy 2461 AGAATAATATTTCTGTTTGTATTTCAATTGGAAGAGCTCTTTGTATTTCACTTGTCT 2520
Db 2461 AGAATAATATTTCTGTTTGTATTTCAATTGGAAGAGCTCTTTGTATTTCACTTGTCT 2520
Qy 2521 TTTCTTTTCTTTTGGCGGATGTTTGAACATGGGGTGTGGCCGCTCAATTCC 2580
Db 2521 TTTCTTTTCTTTTGGCGGATGTTTGAACATGGGGTGTGGCCGCTCAATTCC 2580
Qy 2581 AGTTAAATATGCTGACTTTTGTTCATTTTCAAAAAA 2625
Db 2581 AGTTAAATATGCTGACTTTTGTTCATTTTCAAAAAA 2625

RESULT 2

US-10-424-599-115562

Sequence 115562, Application US/10424599

Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kowalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 115562

LENGTH: 5069

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT3847_75363C.1

US-10-424-599-115562

Query Match 61.1%; Score 1602.6; DB 13; Length 5069;

Best Local Similarity 79.1%; Pred. No. 0;

Matches 1918; Conservative 0; Mismatches 505; Indels 3; Gaps 1;

Qy 1 ATGGGTGAGCGTCTCTCACTGCGCTCCACAGTCTCCGTGAGCGTTTGATAGACCTT 60
Db 500 ATGGGCACCGATCGTTTGAACCGGGTTCACAGTCTCCGTGAGCGCTTGATAGAACCTTC 559
Qy 61 CTGCTCAGAGAACGAGATTTTGGCTTGTCTCAAGATGAGGCAAGAAAGAA 120
Db 560 ACTGCAACAGGAATGAAATTTTGGCTTGTCTCAAGATGAGGCAAGGCAAGG 619
Qy 121 ATTCTGCAACACCATCAATATTCTAGAGTTTGAAGCTATCTCTGAAGAAACAGAAAG 180
Db 620 ATCTGCAACACCAACAGGTCAATCTGATTTGAGAAATCCCTGAGAGAAACAGAG 679
Qy 181 AAGCTCGCTAATGAGCAATTTTGAAGATGAGGCTAGTCAGAGAGCATGCTGTG 240
Db 660 AAGCTCACTGATGCTGCTTTGGAAGATCTTGAATCTACACAGAGAGCATGATTTTG 739
Qy 241 CTTCAATGGGTGCACTTGTCTGTCACAGAGCGCTGTGTTGGAGATCACTTGAAGTG 300
Db 241 CTTCAATGGGTGCACTTGTCTGTCACAGAGCGCTGTGTTGGAGATCACTTGAAGTG 300

Db 740 CCACATGAGGTGCTCTGCTGTCTGTCACAGACCTGCTGTGTGGAGATCCTGAGAGTG 799
Qy 301 AATGTACAGCCCTTGTGTTGAGAACTCACTGTGCTGATATCTTCACTTCAAGAA 360
Db 800 AATGTACAGCCCTTGTGTTGAGAGAGTGGACACTGTGATGACTTGCATTTCAAGAA 859
Qy 361 GAGCTTGTGAGAAAGTTCAAATGAACTTTGTTGAAATTTGAAATTTGAGCCCTTC 420
Db 860 GACTTGTGAGAAAGTTCAAATGAACTTTGTTGAAATTTGAAATTTGAAATTTGAA 919
Qy 421 AACTATCATTTCCCGCCCACTCTTCAAAATCATGTTAATGTTGAGAGTTCTTA 480
Db 920 AATGACAGCTTCCCGCCCACTCTTCAAAATCATGTTAATGTTGAGAGTTCTTA 979
Qy 481 AATGACAGCTTCCCGCCCACTCTTCAAAATCATGTTAATGTTGAGAGTTCTTA 540
Db 980 AATGACAGCTTCCCGCCCACTCTTCAAAATCATGTTAATGTTGAGAGTTCTTA 1039
Qy 541 TTCTCAGAGTCAATGTCACAGAGGCAAGATGTTGAAATGACAGAAATTCAGAAC 600
Db 1040 TTCTCAGAGTCAATGTCACAGAGGCAAGATGTTGAAATGACAGAAATTCAGAAC 1099
Qy 601 TTGAATGCTTTCACATGTTTGAAGAGACAGAGATCTTGTGTAACCTTACTCT 660
Db 1100 TTGAATGCTTTCACATGTTTGAAGAGACAGAGATCTTGTGTAACCTTACTCT 1159
Qy 661 GAGACACATGTCGCAATTTGAAACACCGTTCAGAGAAATCGTTTGAAGAGAGTTG 720
Db 1160 GAGACACATGTCGCAATTTGAAACACCGTTCAGAGAAATCGTTTGAAGAGAGTTG 1219
Qy 721 GATGACACCGGAGAGAGGCTCTGAGATGATTCACACTCTTGTGATCTTCTGAGGCA 780
Db 1220 GATGACACCGGAGAGAGGCTCTGAGATGATTCACACTCTTGTGATCTTCTGAGG 1279
Qy 781 ACTGATCTTCCACCTTGAAGAGTCTTGAAGATCCCATGATGTTCAATGTTGTG 840
Db 1280 ACTGATCTTCCACCTTGAAGAGTCTTGAAGATCCCATGATGTTCAATGTTGTG 1339
Qy 841 ATTCTCACTCCCGACGATTTCTGCTCAAGACATGTTTGGGATATCCGACACCGGT 900
Db 1340 ATTCTCTCCCGACGATTTCTGCTCAAGACATGTTTGGGATATCCGACACCGGT 1399
Qy 901 GGGCAGGTTGTTTCAATCTGATCAAGTCCGAGCTTGAAGATGAGATCTCTCCGT 960
Db 1400 GGGCAGGTTGTTTCAATCTGATCAAGTCCGAGCTTGAAGATGAGATCTCTCCGT 1459
Qy 961 ATTAAGCAACAGGATCTCAACATCAACCCCTGAAATCTCATTTACTAGACTTCTCT 1020
Db 1460 ATTAAGCAACAGGATCTCAACATCAACCCCTGAAATCTCATTTACTAGACTTCTCT 1519
Qy 1021 GATGCTGTGGAACAACTGCGCTCAACGATTTGAGAAAGTATACGAAACAGAGACTCG 1080
Db 1520 GATGCTGTGGAACAACTGCGCTCAACGATTTGAGAAAGTATACGAAACAGAGACTCG 1579
Qy 1081 GATATTTCTTGAAGTACCTTCAAGAAAGAAATGTTTCAAAATGATCTCAAGA 1140
Db 1580 GATATTTCTTGAAGTACCTTCAAGAAAGAAATGTTTCAAAATGATCTCAAGA 1639
Qy 1141 TTTGAAAAAGTCTGGCAATCTTGAACCTTACACAGAGATGTTGCTCATGAAATCTCC 1200
Db 1640 TTTG----AAGTCTGGCCCTTACTTGAACCTTACACAGAGATGTTGCTCATGAAATCTCC 1696
Qy 1201 AAAGAGTTGACAGGACCGCATCTGATCATCGGAAACAAACAGGACGCAATATCTCT 1260
Db 1697 AAAGAGTTGACAGGACCGCATCTGATCATCGGAAACAAACAGGACGCAATATCTCT 1756
Qy 1261 GCTCTCTTGTGCAATAAATGAGTGTCAACAGTGCACCATCGCCATGCTTTGGAG 1320
Db 1757 GCTCTCTTGTGCAATAAATGAGTGTCAACAGTGTCACTAGTGTCACTTGTCAAGCACTT 1816
Qy 1321 AAGACAAATATTCAGATTCAGATATCTATTGAAAGAGCTTGAAGCAATATTCATTTG 1380
Db 1817 AAGACAAATATTCAGATTCAGATATCTATTGAAAGAGCTTGAAGCAATATTCATTTG 1876

Qy	1381	TCCTGCCAAATTACAGCTGATCTTTTTCGATGAAACGTAACAGATTTCATCATCAACGAGT	1440
Dp	1877	TCCTGCCAAATTCACAGCTGATCTTAATTTCGATGAAACGATGAAATTCATTAATCAACGAGT	1936
Qy	1441	ACTTTCACGAGAAATTCGAGGAGGAGGACCTGTGTGTAATACGAGAGCCACACTGCT	1500
Dp	1937	ACCTTCCAGAGATGTCTGGAAGGACGACCTGTGGAACAGTACGAAATCTCAACAGCC	1996
Qy	1501	TTCACTCTCTCTGTGCTCTACCGGTGTGTAAATGTAATCGAATGTGTGATCCCAATTC	1560
Dp	1997	TTCACTCTCTCTGTGCTCTACCGGTGTGTAAATGTAATGTGTGATCCCAATTC	2056
Qy	1561	AACATTTGTTTCCCCGGGTGCTGATATGAGATATATCTTCCCTTACACCGAGAGAAAGCGG	1620
Dp	2057	AACATTTGCTCCCCGGGTGCTGATCAACATTTACTTCCCCACCTGAAACCAAGCCGT	2116
Qy	1621	AGTTGGAAGCATTTCCATCTTGAGATGGAAGACCTTCTTTACACCAAGTTGAGATGAA	1680
Dp	2117	AGTTGGAATCTCTTCCACCTTGAAATCGAAGAACTCTTTACAGTCACTGAGAAATGAA	2176
Qy	1681	GAAACCTAATGTGTGCTCAATGACCGGCAACAAGCCAAATTCGTTCACATGCGCAAGCTT	1740
Dp	2177	GAAACCAATATGTGTGCTGAAAGACCGGCAAGCCAAATTTACTTACCATGGAAGTTG	2236
Qy	1741	GATCGTGTCAAGAACTTAAACGGACCTCTGAGTGTGTCGGGCAAGAACCCAAAGTTGCT	1800
Dp	2237	GATCGAGTGAAGAACATCAACAGACTTGTGAGTGTGTACGGTAAAGACGAACTGAG	2296
Qy	1801	GAGTTGCTAACTCTGTAGTGTGTAGTGTGTATAGCGAAGAACTTAAAGATTTGGAA	1860
Dp	2297	GAGCTGGTGAACCTTGTGTGTGTGTGTGTGGAACAGAGGAAGAGTCAAGACTTGGAA	2356
Qy	1861	GAGAAAGCTGAAATGGAAGAAATGTTGAGCTGATCCGACAGTCAACTTGAACGGCCAA	1920
Dp	2357	GAAAGGCGGAGATGAAGAAATGTAACGGCTGATCGACTTCAAGTTGAACGGCCAA	2416
Qy	1921	TTCCAGATGGAATATCATCTCAATGCAACGAATCCGAATGTGGAATTTGACCTTACCGATACCT	1980
Dp	2417	TTCCAGATGGAATTTATATGCAATGAAACCGTGTGAGGAATGGAAGAGCTTACCGGTATC	2476
Qy	1981	TGCGACACGAAAGGTGCTTTGTATACGCTGTCAATTTATGGAAGCTTTGATGATCAAGTT	2040
Dp	2477	TGCGACACCAAGGGGTGCTTTGTGTGAGCTGTCTATATCGAGGCTTTGTGTGACAGTG	2536
Qy	2041	GTGAGGCAATGACTGTGGGTTTGCAACTTCGCACTGTAAACGGTGAACGACCGAG	2100
Dp	2537	GTGAGGCAATGACTGTGGGCTTGCAACTTCGCAATCAATGCAATGAGTGTGCTGTGAG	2596
Qy	2101	ATTATTTGTCATGGGAAATCTGTGTTTCAACATTTGATCTTACCAATGTGATCAAGTCT	2160
Dp	2597	ATCATTTGTGACCGGAAGTCTGGCTTCAKATTTGACCTTACATGTGACCGTGTGCT	2656
Qy	2161	GACATTACTCGTGCATTTTCTTGAAGAGTGTAAAGAAAGATCACTCACTGGGAATAGATC	2220
Dp	2657	GATCTCTCTTTGTGACTTCTTTGAGAGGTGCAAGCTTGACCCAACTCACTGGGAAGATC	2716
Qy	2221	TCCCAAGGAGCTTGAAGCAATGAGGAGAAAGTATCAATGAAGATTTAATCGGAAGGA	2280
Dp	2717	TCAAGGCTGTGATCTTCACGCGTATTTGAGAGAAAGTACACATGCAAAATTTACTTCAAGG	2776
Qy	2281	CTATTGACCCTCGACGAGAGTATAGATTCTGGAAGCATGTTTCCAACTTTGAACGCGCT	2340
Dp	2777	CTTCTCACTCTCAACGCGTGTCTATAGGCTCTTGAAGAGATGTCTTAACCTTTGAACGCGCT	2836
Qy	2341	GAGAGTGTGCTTACCTTGAAGATTTTATGCTCTTAAATGACCGTAAGCTGTGGAATCA	2400
Dp	2837	GAGAGCGCGCTCTATCTTCGAGATGTTTATGCTCTCAAGTACCGCAAAATGGCTGAGTCT	2896
Qy	2401	GTTCCATTGCGAGAGAGTAATGGA	2426
Dp	2897	GTGCCCCCTGTGCTGAGTAACTGA	2922

```

RESULT 3
US-10-425-114-9604
; Sequence 9604, Application US/10425114
; Publication No. US2004003488B1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 9604
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700852943_FLI
US-10-425-114-9604

Query Match          60.6%; Score 1591.6; DB 13; Length 2601;
Best Local Similarity 78.7%; Pred. No. 0; Mismatches 515; Indels 3; Gaps 1;
Matches 1913; Conservative 0;

QY      1  ATGGCTGAGCGTGTCTCTCACTCGCGCTCCACAGTCTCCGTGAGCGTTGGATGAGACCTT 60
DB      40  ATGGCACTGATGCTGTTGACCCGGGTTCAAGTCTCCGTGAGAGGCTTATGAAACCTTC 99

QY      61  CTGTGCTACAGAAAGAGATTTTGGCCTTGTCTCTAAGATGTAGGGCCAAAGAAAGAA 120
DB      100  ACTGCAACAGAAAGAAATTTTGGCCCTTCTGTCAAGATGAAAGCTTAAAGGCAAGGG 159

QY      121  ATTGCGAAGCAACATCAATATTATCTAGTTTGAAGCTATCCCTGAAGAAAGCAAGAA 180
DB      160  ATCTGCAACACCACAGGTCACTGCTAGTTTGAAGAAATCCCTGAGGAAGAACAGCAG 219

QY      181  AAGCTCGTAAATGTGTCAATTTTGTGAAGTATGAAGGCTAGTCAGAAAGCGATCGTGT 240
DB      220  AAGCTTACTGATGAGGCTTTGGAGAAATCTTGAGATCTACACAGAAAGCCATGATTTG 279

QY      241  CCTCATAGGGTTGCACTTGTCTGTCTCCAAAGGCTCGTGTGGAGATCAATAGAGTG 300
DB      280  CCACCATAGGGTGTCTGTGGCTGTCTCCAAAGGCTGTGTGTGGAGATCACTGAAGTG 339

QY      301  AATGTTCACGCGCTGTGTGTGTGAGGAATCTCACTGTGTGATATCTCCACTTCAAGAA 360
DB      340  AATGTGCAAGCTCTTGTGTGTGTGAGGAGTTGCAACTGCTGATCACTTCAAGAA 399

QY      361  GAGCTTGTGAGGAAGTTCAATATGAAATCTTTGTTTGAATTGGAATTTGAGCCCTTC 420
DB      400  GAACTTGTGATGAAAGTTCTAATGCAAACTTTGTGCTTGAATTTGAAACATTC 459

QY      421  AACTCATATTCCTCCCGCCCAATCTTCAAAATTCGAATTTGTAATAGTGTGAGATTCCT 480
DB      460  AATCAGAGCTTCTCCGCGCCCACTTTAAAGTCAATTTGAAATAGTGTGTGAGTTCCTC 519

QY      481  AATGTCACTTTTGGGCAAAATTTGTCATGACAGAGAGCATGACCCCTTGTGCGAA 540
DB      520  AATGCGCACTTTTCTGCGCAACTCTTCCACGACAGAGAGAGCTTGCACTTTTGGAA 579

QY      541  TTCTCTCAGAGTCAATGTCACAAGGGCAAGAAATGATGTGTAATGACAGAAATTCAGAA 600
DB      580  TTCTCTCAGAGCTTCAAGGCTCAAGGCAAGAAAGCTTTGATGTTGAAATGACAGAAATTCAGAA 639

QY      601  TTGAATGTCTTCAACATGTTTGTGAGGAAGAGCAAGAGATATCTTGTAACCTACTCT 660
DB      640  TCAATGATCACTCAACATGTTCTGAGGAAGAGCTGAGAGATCTGTGGCAAGATGCTCTCT 699

```

QY 661 GAGACACCATGTCGCAATTCGAAACACCGGTTCCAGAAATCGGTTTGAAGAAGGTTGG 720
 DB 700 GAATCCCTCTACTGAAATTTGAGACAAAGTTCCAGAGATTTGGTTTGAAGAGAGGTTGG 759
 QY 721 GGTGACACCGAGAAACGCGTGTCTGAGATGATCCAACTCTTTTGGATCTTCTTGAAGCA 780
 DB 760 GGTGACAAACGAGAGCGTGTCTTGAATCAATTCCTTCTTGGATCTTCTTGAAGGCC 819
 QY 781 ACTGATCTTTCGACCTTTCGAGAGTTCTTTCGAGAAATCCCATGCTTTCATGCTTTCG 840
 DB 820 CCGACCTTCGACCTTTCGAGAGTTCTTTCGAGAAATCCCATGCTTTCATGCTTTCG 879
 QY 841 ATTTCATCTCCCAACGAGTACTTCGCTCAAGCAATGTTTGGGGTATCCGACACCGGT 900
 DB 880 ATTCTTCTTCCCAACGAGTACTTCGCTCAAGCAATGTTTGGGGTATCCGACACCGGT 939
 QY 901 GGGCAGGTTGTTTACATCTTGGATCAAGTCCGAGCTTTTGGAGATGAGATGCTCTCCGT 960
 DB 940 GGGCAGGTTGTTTACATCTTGGATCAAGTCCGAGCTTTTGGAGATGAGATGCTCTCCGT 999
 QY 961 ATTAAGCAACAGAGACTCAATCACTACCCCTGGAATCTCTATTAATAGACTTCTTCT 1020
 DB 1000 ATTAAAGCAACAGAGACTCAATCACTACCCCTGGAATCTCTATTAATAGACTTCTTCT 1059
 QY 1021 GATGCTGTGGAACAAACATGCGGTCAAGCACTTGAAGATGATAGAAACAGAGCACTCG 1080
 DB 1060 GATGCAATGGAACAACTTGTGCGCAAGCTTGAAGAGTGTTCGAAACGAGCACTTC 1119
 QY 1081 GATATTTCTTCGAGTACCTCTTCAGAACAGAAAGGATGTTTCGAAATGATCTCAAGA 1140
 DB 1120 CACATTTCTTCGAGTACCTCTTCAGAACAGAAAGGATGTTTCGAGATCTCAAGA 1179
 QY 1141 TTTGAAAAAGTCTGCGCATCTTGGAAACCTTCAAGAGATGTTCTCAATGAATCTTC 1200
 DB 1180 TTCCG---AAGTCTGCGCATCTTGGAAACCTTCAAGAGATGTTCTCAATGAATCTTC 1236
 QY 1201 AAAGAGTTGACGCGACGCGCATCTGATCATTCGGAACCAAGAGACGCGCAATTCGTC 1260
 DB 1237 AAAGAGTTGACGCGACGCGCATCTGATCATTCGGAACCAAGAGATGAGAAACATTTGT 1296
 QY 1261 GCTCTCTTGTCTGCAATTAATTAAGTGTACACAGTGCACCATGCTCCATGCTTGGAG 1320
 DB 1297 GCTCTTGTGTGGCATTAATTAAGAGTCACTGATGACATTTGCTCATGCACTTGAAG 1356
 QY 1321 AAGACAAATATCCGATTCAGATATCTAATGGAAGAACTTGAAGCAATATCAATTTT 1380
 DB 1357 AAGACAAATATCCCGAATCCGACATTTAATGGAAGAAATGGAAGAGATACCACTTC 1416
 QY 1381 TCTTGCCAAATTTAAGCTGATCTTTTTCGAATGAACCAATAGATTTTCATCATCAAGT 1440
 DB 1417 TCTTGCCAAATTTAAGCTGATCTTTTTCGAATGAACCAATAGATTTTCATCATCAAGT 1476
 QY 1441 ACTTTCAGAGAAATTCAGAGAGAGCACTGTTGTGATATACAGAGCACAACCTGCT 1500
 DB 1477 AACTTTCAGAGAGATGCTGGAAGAGCACTGTTGTGATATAGAGTCTCACACACCC 1536
 QY 1501 TTTACTCTTCTGCTCTACCGTGTGTAATGATGATGATGTTTGAATCCCAATTC 1560
 DB 1537 TTTACTCTTCTGCTCTACCGTGTGTAATGATGATGATGTTTGAATCCCAATTC 1596
 QY 1561 AAGATGTTTCCCGGAGTGTGATGAGATATATCTTCCCTTACACGGAAGAGAGGG 1620
 DB 1597 AAGATGTTTCCCGGAGTGTGATGAGATATATCTTCCCGGAGAGATGATGAGAGGG 1656
 QY 1621 AGTTGAAGCAATTTTCATCTGAGATGAGAGCACTTCTTTTACCAACCAAGTTGAGATGA 1680
 DB 1657 AGTTGAAGCACTCTTCCACCCGGAATTCGAAGAACTTCTTTACAGCTCTGTGAGAGATGA 1716
 QY 1681 GACACCTTATGTGTCTCAATGACGCAACAGCAATTTGTTTCAATGCAAGGCTT 1740
 DB 1717 GACACATATGTGTCTCAATGACGCAACAGCAATTTATCTTCAATGCAAGGCTT 1776

QY 1741 GATGCTGTCAAGAACTTAACCGGACTGCTGATGCTGTCGCGCAAGAACCCAAAGTTGCT 1800
 DB 1777 GACGCTGTGAAGAACATCAACAGACTGCTGTGAGTGTATCGGTAAAGAACGCAAGCTGAG 1836
 QY 1801 GATGCTGTCAACCTTCGATGTTGTAGTGTATAGGCGAAAGAAATCTTAAGATTTTGA 1860
 DB 1837 GAGTTGTGAACCTTGTGTGTGTGCGGAGACAGAGAGAGAGATGCAAGGACTTGA 1896
 QY 1861 GAGAGGCTGAATGAAGAAATGTTTGAAGCTGATGCAAGTCAACTTGAACGCGCA 1920
 DB 1897 GAGAGGCTGAATGAAGAAATGTTTGAAGCTGATGCAAGTCAACTTGAACGCGCA 1956
 QY 1921 TTCAGATGATATCATCTCAATGAACAAATCCGAAATGTTGAATTTACCGATACAT 1980
 DB 1957 TTCAGATGATATTTATCTCAATGAACAAATCCGATGAGAAACGAGAGCTGTACCGTGTATC 2016
 QY 1981 TCGCACAGAAAGTGTCTTGTATACAGCTGCAATTTGATGAAGCTTGTGATGACAT 2040
 DB 2017 TCGCACACAAAGGAGCTTGTGTACAGCGCTATATACGAGGCTTGTGTGACATGT 2076
 QY 2041 GTGAGGCAATGATCTGCGGTTTGCACAACTTGCACCTGTAAACGCTGACACGCGAG 2100
 DB 2077 GTGAGGCAATGATCTGCGGTTTGCACAACTTGCACCTGTAAACGCTGACACGCGAG 2136
 QY 2101 ATTATTTGATGAGGAAATCTGTTTCAACATGATCTTACATGATGATCAAGTCT 2160
 DB 2137 ATCATTTGATGAGGAAATCTGTTTCAACATGATCTTACATGATGATCAAGTCT 2196
 QY 2161 GACATCTGCTGATTTCTTTGAAAGTGTAAAGATTCATCTCACTGGATTAAGATC 2220
 DB 2197 GATCTCTTGTGATCTTCTTTGAGAGTGAAGCTTGACCCCAACCTGAGAAACATTC 2256
 QY 2221 TCCCAAGGCTTGAACGAATGAGAGATATACATGAGAAATTTACTCGAGAGA 2280
 DB 2257 TCCCAAGGCTTGAACGAATGAGAGATATACATGAGAAATTTACTCGAGAGA 2316
 QY 2281 CTATTTGACCTTGACAGAGATGATGATCTTGAAGCAATGTTTCAACCTTGAACGCGT 2340
 DB 2317 CTTCCTCATCTCATCTGCTGTATGAGCTTCTGGAAGCATGATGATCACTTGAACGCGT 2376
 QY 2341 GAGAGTGTGCTTACCTTGAAGATGTTTATGCTTTAAGTACCTGAGCTGATCA 2400
 DB 2377 GAGAGGCGCGCTTATCTCGAATGTTTATGCTTCAAGTACCCCAATGCGCGAGTCT 2436
 QY 2401 GTTCATTTGACAGAGGATTAATGAACCTG 2431
 DB 2437 GTGCCCTTGTCTGTGATGAATGAGAGAT 2467

RESULT 4
 US-10-424-599-12144
 ; Sequence 12144, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 12144
 ; LENGTH: 3101
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_110974C.1
 US-10-424-599-12144

Query Match 60.6%; Score 1590; DB 13; Length 3101;

Best Local Similarity 78.7%; Pred. No. 0; Matches 1912; Conservative 0; Mismatches 516; Indels 3; Gaps 1;

1 ATGGCTGACGGCTGCTCACTGCGCTCCAGCTCCCGTAGCGCTTGATGAGACCTT 60
131 ATGGCCATGATCGTTGATCCCGGTTCAAGTCTCCGTGAGAGCTTGATGAACTTC 190
61 CTGGCTCAGAGAACGAGATTTTGGCTTCTCAGAGATCGAGGCAAGAAAGGA 120
191 ACTGCCAACAGGAAGAAATTTTGGCCCTTCTGTCAAGATCGAAGCTAAGGGCAAGGG 250
121 ATTCTGCACACCATCAATTAATTCTAGAGTTGAAGCTATCCCTGAAGAGAACGAAG 180
251 ATCTGCACACCAACACAGATCAATGCTGAGTTGGAATCCCTGAGAGAACGAGCAG 310
181 AAGCTGCTAATGAGGCAATTTTGAAGATGAAGGCTAGCAGAGAGCGATCGTTTG 240
311 AAGCTTACTGATGAGTCCCTTGGAGAAAGTCTTGAGATCTACAGAGAAAGCCATAGTTTG 370
241 CCTCATGAGGTTGCACTTGTGTTGCTCCAAAGGCTGAGTTGTTGGAGTACATTAGAGTG 300
371 CCACCATGGGTTGCTCTGGCTGTTGCTCCAAAGGCTGAGTGTGGAGATACCTGAAGTG 430
301 AATGTTACGCGCTTGTGTTGTTGAGAACTCACTGTTGCTGAGTATCTCACTTCAAGGA 360
431 AATGTCACGCTCTGTTGTTGAGAGAGTTGCAACCTGCTGAGTACCTTCACTTCAAGGA 490
361 GAGCTTGTGATGAGAAAGTTCAAACTTTGTTTGAATTGGAATTTTGAAGCCCTTG 420
491 GAACTTGTGATGAGAAAGTTCTAATGCAACTTTGTTGCTGAGTTGCACTTTGAACATTC 550
421 AACTCATGATTTCCCGGCAACTCTTTCAAAATCATTTGTAATGATGTGAGAGTTCTTA 480
551 AATGACAGCTTCCCTCGCCAACTCTTAACAAGTCAATTTGAAATGATGCAAGTTCTCTC 610
481 AATGTCACCTTTGGGCAAAATTTGTTCAATGACAAAGAGAGACATGACCTTTGCTGGA 540
611 AAGCGCACCTTTCTGCGCAAACTTTCCACGACAAAGAGAGCTTGCAACCTTTTGGAA 670
541 TTCTCAGAGTCATTTGTCACAAAGGCAAGAACATGATGTTGAATGACAGAAATTCAGAAC 600
671 TTCTCAGAGCTTCAACAGCTGCAAGGAAAGACTTTGATGTTGAATGACAGAAATTCAGAAC 730
601 TTGAATGCTCTTCAACATGTTTGAAGAAAGAGAGATCTTGTATCCCTACCTCT 660
731 CCAAGTGCATCCAAATGTTCTGAGAAAGCTGAGAGATCTGAGGACAGTGCCTCT 790
661 GAGACACCATGTCGGAATTTGGAACACCGGTTCCAGAAATCGGTTTGGAAAGAGTTG 720
791 GAAACTCCCTACTCAGAAATTTGAGCAAAATTCAGAGATTTGTTGGAAGAGGAGTG 850
721 GATGACACCGGAGAGCGGTGCTGAGATGATCCAACTCTTTTGGATCTTCTTGAAGCA 780
851 GGTGACACCGAGAGCGGTGCTTGAATCAATCTCTCTTGGATCTTCTTGAAGCC 910
781 ACTGATCTTGCACCTTGAAGAGTTCTTGGAGAAATCCCATGAGTTCAATGTTTG 840
911 CCTGACCTTGCACCTTGAAGCTTCTTGGAGAAATTCCTATGATCTTCAATGTTTG 970
841 ATTCTCATCTCCCAAGGATCTTGGCTCAAGACAAATGTTTGGGATATCCGACACCGGT 900
971 ATTCTTCTCCCAAGGATCTTGGCTCAAGATGATGTTGGAATACCTTGAACCTGAGT 1030
901 GGGCAGGTTGTTTACATCTTGGATCAAGTCCGAGCTTGGAGAAATGAGATGCTCCTCGT 960
1031 GGGCAGGTTGTTTACATCTTGGATCAAGTCTGCTGTTGGAGAAAGAGATGCTCATATG 1090
961 ATAAAGCAACAAGATCTCAACATCAACCCCTGGAATCTCATTAATTAATCAATCTTCT 1020
1091 ATTAAGCAACAAGATGGAATGATCTGATCTGATCTCATTAATCAACCCGCTTCTCC 1150
1021 GATGCTGTGGAACAACATGCGGTCAAGCATTTGAGAAAGTATAGGAACAGAGCATCTG 1080
|||||

1151 GATGCAATCGGAACACTTGTGCGCAACGCTTGAGAAAGTGTGCGAACCGAGCATCC 1210
1081 GATATTTCTGAGTACCTTTCAGAAACAGAAATTTGTTCCAAAATGATCTCAAGA 1140
1211 CACATTTCTGAGTATCCCTTTTGAATGAGAAAGGAATTTGTTCTGATGATCTCAAGA 1270
1141 TTGAAAAGTGTGCGCAATCTTGGAAAAGCTTACACAGAGATGTTGCTCATGAATCTCC 1200
1271 TTCC---AAGTGTGCAATCTTGGAAAAGCTTACAGAGATGTTGCTCATGACCTTGC 1327
1201 AAAAGTTGACCGGACCGCAATCTGATGATCGGAACNACAGCGAATATGCTC 1260
1328 AAAAGTTGACCGGACCGCAATCTGATGATCGGAACNACAGCGAATATGCTC 1387
1261 GCTCTCTTGTGCGCAATTAATTTAGTGTCAACAGTGCACCATGCGCATGTTGGAG 1320
1388 GCTCTTGTGCGCAATTAATTTAGTGTCAACAGTGTGACATGCTCATGACCTTGGAG 1447
1321 AAGCAAAATATCCAGATTCAGATATCTAATGGAAGAGCTTGAAGCAAAATACATTC 1380
1448 AAGCAAAATATCCCGAATCCGACATTTAATGGAAGAGATATCACTTC 1507
1381 TCTTGCAATTTAAGCTGATCTTTTGGCAATGAACATACAGTTTCAATGATCAACAGT 1440
1508 TCTTGCAATTTAAGCTGATCTTTTGGCAATGAACATACAGTTTCAATGATCAACAGT 1567
1441 ACTTTCAGGAATTTGAG 1500
1568 ACTTTCAGGAATTTGAG 1627
1501 TTCACTTCTGCTGCTCTCAACGTTGTTGATCAGTATCGATGTTGATGATCCAAATTC 1560
1628 TTTCCTTCTGCTGCTCTCAACGTTGTTGATCAGTATCGATGTTGATGATCCAAATTC 1687
1561 AACTGTTTCCCTGAGTATGAGATGAGATATCTTCCCTTACACCGAAGAGAGAGAG 1620
1688 AACTGTTTCCCTGAGTATGAGATGAGATATCTTCCCTTACACCGAAGAGAGAGAG 1747
1621 AGTTGGAAGATTTCCATCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
1748 AGTTGGAAGATTTCCATCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1807
1681 GAACATTAATGTTGCTCAATGACCGCAACAGCAATCTGTTTCAATGACCAAGCTT 1740
1808 GAACATTAATGTTGCTCAATGACCGCAACAGCAATCTGTTTCAATGACCAAGCTT 1867
1741 GATGTCGCAAGAACTTAACCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
1868 GATGTCGCAAGAACTTAACCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1927
1801 GATGTCGCAAGAACTTAACCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
1928 GATGTCGCAAGAACTTAACCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1987
1861 GAGAGGCTGAATGAGAAATTTTGAAGTATGATGACAGATGACAACTTGAACGCGCA 1920
1988 GAGAGGCTGAATGAGAAATTTTGAAGTATGATGACAGATGACAACTTGAACGCGCA 2047
1921 TTGAGATGATATGATCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
2048 TTGAGATGATATGATCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2107
1981 TGCGACAGAAAGGTGCTTGTATGACAGCTGATGATGAGAGAGAGAGAGAGAGAGAG 2040
2108 TGCGACAGAAAGGTGCTTGTATGACAGCTGATGATGAGAGAGAGAGAGAGAGAGAG 2167
2041 GTGAGGCAATGATGCTGCTGTTGCAACATTTGCAACCTGTAACGTTGACAGCGAG 2100
2168 GTTGGGCAATGATGCTGCTGTTGCAACATTTGCAACCTGTAACGTTGACAGCGAG 2227
2101 ATTAATGTCATGAGGAATCTGTTTCAACATGATGATGATGATGATGATGATGATGAT 2160
2228 ATCAATGTCATGAGGAATCTGTTTCAACATGATGATGATGATGATGATGATGATGAT 2287
|||||

QY 2161 GACATACGTCGATTTCTTTGAAAGTGTAAAGATTCATCTCATGGGATTAAGTC 2220
DB 2288 GATCTCTCTTTGATCTCTTTGAGAGTGTACAGCTTGAACCCATCGGAGAAACATC 2347
QY 2221 TCCCAAGAGGCTTGAACGATAGAGAGATATACATGAAGATTACTCGAGAGA 2280
DB 2348 TCAAGAGCTGGTCTCCAGCGATTAAGAGAGATACATGCAAAATTTACTCAGAGG 2407
QY 2281 CTATTGACCTGACAGAGATGTATGATTTCTGGAAGCATTTTCCAACTTGAACGCCGT 2340
DB 2408 CTCTCACTCTCACTGGTGTCTATGCTTCTGGAAGCATGTGTCTAACTTGAACGCCGT 2467
QY 2341 GAGAGTGTGTCTCTTTGAGATGTTTATGCTTAACTGACCGTAAAGCTGTAATCA 2400
DB 2468 GAGAGCGCGCTATCTCGAGATGTTCTATGCTCTCAAGTACCGCAATTTGGCCGAGCT 2527
QY 2401 GTTCCATTGGCAGAGAGATTAATTGAACCTG 2431
DB 2528 GTGCCCTCTGCTGTGAGTAACTGAGGATG 2558

RESULT 5
US-10-425-114-6460
Sequence 6460, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21 (5313) B
CURRENT APPLICATION NUMBER: US/10/425.114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 6460
LENGTH: 2772
TYPE: DNA
ORGANISM: Glycine max
FEATURES:
OTHER INFORMATION: Clone ID: 700565776_FLI
US-10-425-114-6460

Query Match Best Local Similarity 59.8%; Score 1570.4; DB 13; Length 2772;
Matches 1912; Conservative 0; Mismatches 527; Indels 6; Gaps 2;

QY 1 ATGCGTGAAGCGTCTCTCACTCGCGTCCAGAGTCTCGTGAAGCGTTGGATGAGACCTT 60
DB 53 ATGCAATATCATCTTTGACACATCTCTCACTCTTCCGAGAGGTTTGAAGAACTCTC 112
QY 61 CTGCTCACAGAGAGAGATTTTGGCTTGTCTCAAGATCGAGGGAAGAAAGGA 120
DB 113 ACTGTCACAGAGAGAGATTTTGGCTTGTCTCAAGGCTTGAAGGCAAGGGA 172
QY 121 ATTCTGCAACACATCAATATTTCTAGAGTTTGAAGTATCCCTGAGAGAAAG 180
DB 173 ATCTGCAACATCAACAGAGTGTGAGAGTTTGAAGAAATCCCTGAGAGAAAG 232
QY 181 AAGCTCGTAATGAGTATTTTGAAGTATTAAGGCTAGTCAAGAGCATCGTGTG 240
DB 233 AAATCTCAAGATGAGTCTTTGAGAAAGTTTGAATCACAAGAGCATAGTGTG 292
QY 241 CCTCAATGAGTGTGATCTTGTCTGTCAGAGCGCTGTGTTTGGAGTATTAAGTGTG 300
DB 293 CCACCATTTGAGCTCTTGTGTCAGCAAGGCTGTGTATGAGAAATCTCGTGTG 352
QY 301 AATGTCACGCGCTTGTGTTGAGAACTCACTGTTGAGTATCTCACTTCAAGGA 360

DB 353 AATGTCACATGCTGTGTGTGATGAGCTTCTGCTGAGTATCTGCTTCAAGAG 412
QY 361 GAGCTGTGATGAGAGTCAATGAGAACTTTGTTGAAATGATTTGAGACCTTC 420
DB 413 GAGCTGTGATGAGAGTCAATGAGAACTTTGTTGAAATGATTTGAGACCTTC 472
QY 421 AACTCATATTCCTCCCGCCCACTCTTCAAAATTCATGTAATGATGAGATTCCTA 480
DB 473 AATGATCTCTTCTCCCGCCCACTCTGACAGATGATGAAATGAGATTCCTC 532
QY 481 AATGTCACCTTTGGCAAAATTTGTTCAATGACAGAGAGATGACCTTTGCTGAA 540
DB 533 AATGTCACCTTTGGCAAAATTTGTTCAATGACAGAGAGATGACCTTTGCTGAA 592
QY 541 TTCTCAAGTCTCAATGTCACAGAGGCAAGATGATGTAATGACAGATTCAGAAC 600
DB 593 TTCTCAAGTCTCAATGTCACAGAGGCAAGATGATGTAATGACAGATTCAGAAC 652
QY 601 TTGATGCTCTTCAACATGTTTGAAGAGAGAGATGATCTGATCCTACCTCT 660
DB 653 CTGATCTCTTCAACATGTTTGAAGAGAGATGATCTGATCCTACCTCT 712
QY 661 GAGACACATGTCGGAATTCGAACACCGGTTCCAGAAATCGTTGAAAGAGTTG 720
DB 713 GAAACACCTTACTCGGAATTCGAAGACAGATTCGAGAGATGAGTGTGAGAGG 772
QY 721 GGTGACACCGGCAAGAGCGGTCTGAGATGATGATCCTCTTTGATCTTTGAGGCA 780
DB 773 GGTGACACCGGCAAGAGCGGTCTGAGATGATGATCCTCTTTGAGGCA 832
QY 781 ACTGATCTCTGACACCTTGAAGAGTCTTGGAGAGATTCCTGATGATGATGATG 840
DB 833 CTGATCTCTGACACCTTGAAGAGTCTTGGAGAGATTCCTGATGATGATGATG 892
QY 841 ATTCTCACTCCCAAGAGATCTGCTCAAGACATGTTTGGAGATTCCTGATGATG 900
DB 893 ATCTCTCTCCCAAGAGATCTGCTCAAGACATGTTTGGAGATTCCTGATGATG 952
QY 901 GGGAGATGTTTGAATCTTGAATGATGATGATGATGATGATGATGATGATGATG 960
DB 953 GGAAGATGTTTGAATCTTGAATGATGATGATGATGATGATGATGATGATGATG 1012
QY 961 ATAAAGCAAGAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1020
DB 1013 ATCAAGAAAGAGATCTTGAATGATGATGATGATGATGATGATGATGATGATG 1072
QY 1021 GATGCTGTGGAACATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
DB 1073 GATGCTGTGGAACATGATGATGATGATGATGATGATGATGATGATGATGATG 1132
QY 1081 GATGCTGTGGAACATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
DB 1133 GATGCTGTGGAACATGATGATGATGATGATGATGATGATGATGATGATGATG 1192
QY 1141 TTTGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
DB 1193 TTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1249
QY 1201 AAGAGTGTGACAGGACGAGATCTGATGATGATGATGATGATGATGATGATGATG 1260
DB 1250 AAGAGTGTGACAGGACGAGATCTGATGATGATGATGATGATGATGATGATGATG 1309
QY 1261 GCTCTCTTGTGACAGGACGAGATCTGATGATGATGATGATGATGATGATGATGATG 1320
DB 1310 GCTCTCTTGTGACAGGACGAGATCTGATGATGATGATGATGATGATGATGATGATG 1369
QY 1321 AAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
DB 1370 AAGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1429
QY 1381 TCTGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
DB 1430 TCTGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1489

1441 ACTTCCAGAAATTGACAGAGACACCTGTTGTAATACGAGAGCCACCTGCT 1500
1490 ACCTTCCAGAGATTGTCGAGAGACACCTGTTGAGATGAGAGTCACACTGCCC 1549
1501 TTCACTCTCTGCTGCTACCGCTGTTACATGCTATCGATGCTTGAATCCAAATTC 1560
1550 TTCACTCTCTGAGACCTTACCGCTGTTACATGCTATGATGCTTGAATCCAAATTC 1609
1561 AACATGTTTCCCTGCTGCTGATATGAGATATATCTTCCCTTACACCGAAGAGCGG 1620
1610 AACATGCTCTCTCCCGTGCAGACATGAGTATATCTTCCATACCTGAAACCTGAGGT 1669
1621 AGTTGAAGCATTTCCATCTGAGATCGAAGACCTTCTTTACACCAAGTTGAGATGAA 1680
1670 AGTTGAAGCATTTCCATCTGAGATCGAAGACCTTCTTTACAGCTCAGTGAAGATGAG 1729
1681 GAACACTATGCTGCTCAATGACCGGCAACAGCCAAATCTGTTTCAATGCCAAGGCTT 1740
1730 GAACACTATGCTGCTCAATGACCGGCAACAGCCAAATCTTCCATGCTGCAAGGCTT 1789
1741 GATCGTCTCAAGAACTTAACCGGACCTGCTGAGTGTGCGGCAAGAACCCAAAGTGGCT 1800
1790 GACCTGTGAAGAAACATCAACGGGCTTGTGAGTGTATCGGGAAGAACGACGCTCCGC 1849
1801 GAGTTGCTAACTCTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 1860
1850 GAGTTGCTAACTCTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 1909
1861 GAGAAAGCTGAAGAAAGAAATGTTTGAAGCTGATCGCAAGTCAATTTGAACGCGCAA 1920
1910 GAGAAAGCTGAAGAAAGAAATGTTTGAAGCTGATCGCAAGTCAATTTGAACGCGCAA 1969
1921 TTCAATGATGATATCTCAATGAAACAGAAATCCGAATGTTGAACCTTACCGATACAT 1980
1970 TTCAATGATGATATCTCTCAATGAAACAGAAATCCGAATGTTGAACCTTACCGATAC 2029
1981 TGCGACAGAAAGGCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2040
2030 TGCGACAGAAAGGCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2089
2041 GTGAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
2090 GTGAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2149
2101 ATTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
2150 ATCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2209
2161 GACATATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
2210 GAGATCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2269
2221 TCCCAAGAGAGCTGAAACGAAATGAGAGAAATGAGAGAAATGAGAGAAATGAGAGAAATG 2280
2270 TCCCAAGAGAGCTGAAACGAAATGAGAGAAATGAGAGAAATGAGAGAAATGAGAGAAATG 2329
2281 CTATGACCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
2330 CTCTGACCTGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2389
2341 GAGAGTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
2390 GAGAGCAACCTTACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2449
2401 GTTCCATTGGC---AGAGAGTAAATGAACTGTTAAATTAACAT 2442
2450 GTGCCCTTGTCTATGAGAGTAAATGATGATGATGATGATGATGATGATGATGATGATG 2494

Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ. ID NOS: 285684
; SEQ. ID NO 11458
; LENGTH: 3360
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110354C.1
US-10-424-599-11458

Query Match 59.8%; Score 1570.4; DB 13; Length 3360;
Best Local Similarity 78.2%; Pred. No. 0;
Matches 1912; Conservative 0; Mismatches 527; Indels 6; Gaps 2;

1 ATGCTGAGGCTGCTCACTCGGCTCACAAGTCTCCGTGAGCGTTGGATGAGACCTT 60
172 ATGCAATATCATCTTGTGACACACTCTCACTTTCGCGAGAGTTGATGAAATCTTC 231
61 CTGCTGACAGAAAGAGATTTTGGCTGCTGCTCAAGATCGAGGCAAGAAAGAA 120
232 ACTGCTACAGAAAGAAATTTTGGCTGCTGCTCAAGATCGAGGCAAGAAAGAA 291
121 ATTCTGACACCATCAATTAATTTAGATTTGAAGCTATCCCTGAAGAAACAGAA 180
292 ATCTGCAATATCAACAAAGGTTGAGAGTTGAAATCCCTGAGAGACAGAAAG 351
181 AAGTCTGATATGATGATTTTGAAGTATGAAAGCTAGTCAAGAAAGCATGTGTG 240
352 AATCTCAAGATGATGATTTTGAAGTATGAAAGCTAGTCAACAGAAAGCATGTGTG 411
241 CCTCATGAGGTTGATGATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
412 CCACATTTTGTAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 471
301 AATGTCACGCTTGTGTTGAGAACTCACTGTTGCTGATATCTCACTTCAAGAA 360
472 AATGTCACAGTGTGTTGATGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 531
361 GAGCTTGTGATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 420
532 GAGCTTGTGATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 591
421 AACTCATATTTCCCGGCAACTCTTCAAAATCCATTGATATGATGATGATGATGATG 480
592 AATGTCATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
481 AATGTCACCTTTCGCAAAATTTTTCATGACAGAGAGATGCAACCTTTCGCTGAA 540
652 AACGCACTTTCGCAAGCTCTTTCATGACAGAGAGATGCAACCTTTCGCTGAA 711
541 TTCTGAGATGCTTTCATGACAGAGAGATGATGATGATGATGATGATGATGATGATG 600
712 TTCTGAGATGCTTTCATGACAGAGAGATGATGATGATGATGATGATGATGATGATG 771
601 TTGAATGCTTTCATGACAGAGAGATGATGATGATGATGATGATGATGATGATGATG 660
772 CTGATATCTTTCATGACAGAGATGATGATGATGATGATGATGATGATGATGATGATG 831
661 GAGACACCATGCTGCAATGAAACCGGTTCCAGAAATCGTTTGAAGAGTGTG 720
832 GAAACACCTTACTGCAATGAAACCGGTTCCAGAAATCGTTTGAAGAGTGTG 780
721 GTGACACCGGCAAGCGGTTCTGAGATGATGATGATGATGATGATGATGATGATGATG 780

```

Db      892  GGTGACACTGCGAGCGGTGCTCGAGATGATCCAGCTTCTCGGACCTTCTTGAGGCA 951
Qy      781  ACTGATCTTGCACCCCTTGAGAGTTCTTGGAGAAATCCCAATGGTGTCAATGTGTG 840
Db      952  CCTGACCCCTTGACCCCTTGAGACATCTCTTGGAAAGTCCCTATGTCTTCAATGTGT 1011
Qy      841  ATTGCACTCCCAAGATCTGCTGCTGAGACATGTTTGGGGTATCCGACACCGGT 900
Db      1012  ATCTTTCTCCCAAGTACTTGGCCAGATATGTTCTTGGATACCTTGACCTGGT 1071
Qy      901  GGGGAGGTTGTTTATCATCTTGGATCAATCGAGCTTGGAGAAATGAGATCTCTCGT 960
Db      1072  GGACAGGTTGTTTATCATCTTGGATCAATGTTGCTTGGAGAAATGAGATCTCAACGC 1131
Qy      961  ATAAAGCAACAGATCAATCACTACCCCTGGAATCTCTATTTACTATGACTTCTCT 1020
Db      1132  ATCAAGAAACAAAGGCTTGATATCACTCCCTGTATCTCATTTATCACTGCTCTCT 1191
Qy      1021  GATGCTGCGAACAACATGGGGTCAAGACTGAGAAATGATCGAAGACAGACCTCG 1080
Db      1192  GATGAGTAGAATCTACCTGTGGCCACGCTGAGAGGGTATGATATCTGAATATGT 1251
Qy      1081  GATATTTCTGAGTACCTTCAAGACAGAAAGGGAATGTTCCGAAATGATCTCAGA 1140
Db      1252  GACATTTCTCAGATTCCTTTCAGAACAGAAAGGGAATGTTCCGAAATGATCTCAGA 1311
Qy      1141  TTTGAAAAAGTCTGGCCATCTTGGAAATCTTACACAGAGATGTTCTCATGAATCTCC 1200
Db      1312  TTCC---AGTCTGGCCATCTTACAGAGACTTACACTGAGAGATGTCCTTGAACCTG 1368
Qy      1201  AAAGAGTTCAGGAGCGGAGATCTGATCTGATCGGAAACAGGAGCGGAATATGCTC 1260
Db      1369  AAGAGTTCAGGAGCGGAGATCTGATCTGATCGGAAACAGGAGCGGAATATGCTC 1428
Qy      1261  GCCTCTTGTCTGCAATAAATGAGTGTCAACAGTGCACCATCGCCCATGCTTTGAG 1320
Db      1429  GCCTCTTGTGTAGCAATAAATGAGATTAATCTGATGATCAATGCTCATGCTCTAGA 1488
Qy      1321  AAGACAAATATCCAGATTCAGATATCTATTTGAAAGAACTTGAAGACAAATACATTC 1380
Db      1489  AAGACAAATATCCAGATTCAGATATCTATTTGAAAGAAATATCATATTC 1548
Qy      1381  TCTTGCCAAATTAAGCTGATCTTTTGCATGAAACCATATGATATCATCAACGAT 1440
Db      1549  TCATGCCAAATTAAGCTGATCTTTTGCATGAAACCATATGATATCATCAACGAT 1608
Qy      1441  ACTTTCAGAAATTTGAGAGAGCACTGTTGTCAATACGAGGCCACATGCT 1500
Db      1609  ACCTTTCAGAGATTTGCTGAGAGAGCACTGTTGAGACATGATGAGACATGCTGCC 1668
Qy      1501  TTCACTCTTCTGCTCTGATCCGTTGTATCAATGATGATGATGATGATCCAAATTC 1560
Db      1669  TTCACTCTTCTGATCCGTTGTATCAATGATGATGATGATGATGATGATCCAAATTC 1728
Qy      1561  AACATTTGTTTCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1620
Db      1729  AACATTTGTTTCCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1788
Qy      1621  AGTTTGAAGATTTTCCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db      1789  AGTTTGAAGATTTTCCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1848
Qy      1849  GAAACATATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1908
Db      1941  GATGCTGCAAGAACTTAAACCGGAGCTGCTGAGTGTGCTGAGAGAAACCGGAGCT 1800
Qy      1909  GACGCTGCAAGAACTTAAACCGGAGCTGCTGAGTGTGCTGAGAGAAACCGGAGCT 1968
Db      1801  GAGTGTGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860

```

```

Db      1969  GAGTGTGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2028
Qy      1861  GAGAGGCTGAATTAAGAAATGTTTGAAGTATGATGATGATGATGATGATGATGAT 1920
Db      2029  GAGAGGCTGAATTAAGAAATGTTTGAAGTATGATGATGATGATGATGATGATGAT 2088
Qy      1921  TTTGATGATATATCTCAATTAAGAAATGATGATGATGATGATGATGATGATGAT 1980
Db      2089  TTTGATGATATATCTCAATTAAGAAATGATGATGATGATGATGATGATGATGAT 2148
Qy      1981  TGGACACGAAAGGCTGTTTGAACAGCTGATGATGATGATGATGATGATGATGAT 2040
Db      2149  TGTACACAAAGGCTGTTTGAACAGCTGATGATGATGATGATGATGATGATGAT 2208
Qy      2041  GTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Db      2209  GTTGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2268
Qy      2101  ATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db      2269  ATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2328
Qy      2161  GACATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Db      2329  GAGATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2388
Qy      2221  TCCCAAGAGGCTTGAAGCAATGATGATGATGATGATGATGATGATGATGATGAT 2280
Db      2389  TCCCAAGAGGCTTGAAGCAATGATGATGATGATGATGATGATGATGATGATGAT 2448
Qy      2281  CTATTTGACCTGACAGAGTATGATGATGATGATGATGATGATGATGATGATGAT 2340
Db      2449  CTCTTTGACCTGACAGAGTATGATGATGATGATGATGATGATGATGATGATGAT 2508
Qy      2341  GAGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
Db      2509  GAGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2568
Qy      2401  GTTCCATTTGCT---AGAGGATTAATGAAACCTGTTAAATTAACAT 2442
Db      2569  GTTCCATTTGCT---AGAGGATTAATGAAACCTGTTAAATTAACAT 2613

RESULT 7
US-10-424-599-11459
; Sequence 11459, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 11459
; LENGTH: 3438
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110355C.1
US-10-424-599-11459

Query Match          59.7%; Score 1568; DB 13; Length 3438;
Best Local Similarity 77.9%; Pred. No. 0;
Matches 1902; Conserved 0; Mismatches 536; Indels 3; Gaps 1;

Qy      1  ATGCTGAGGCTGCTCACTGCGTCCAGAGTCTCCGTGAGCGTTGATGAGACCTT 60
Db      689  ATGCAATATCACTTTGACACACTCTCACTCTTTCGCGAGAGGTTGATGAATCTCTC 748

```

OY	61	CTTGCTCAGAGAAAGAAATTTTGGCCTTGCTCTCAAGATCGAGGGCAAGAAAAGA	120
Db	749	ACTGCTCACAGGAAGAAATTTTGGCCCTTTTGTAAAGCTTGAAGCCAAAGGCAAGGA	808
OY	121	ATTCTGCAACACCATCAATTAATTTTAAAGTTTGAAGCTATCCCTGAAGACAAGAAG	180
Db	809	ATCTCGACAACACCAAGGTGGTTCAGAGTTTGAAGAAATCCCTGAAGACAGAAAG	868
OY	181	AAAGCTCGTAATGTGCATTTTTTGAAGTATTTGAAGGCTAGTCAGAAAGCATCGTGTG	240
Db	869	AAAGCTCAAGGGTGTGCTTTTGAAGAAATTTTGAATCTACACAGGAAGCATAGTGTCTG	928
OY	241	CCTCCATGGGTTGCACTTGCTGTGCTCCAAAGCCCTGGATTTGGGAATCAATTAAGTCTG	300
Db	929	CCACCATTTTGTGGCTCTGGCTGTTTGACCAAGCCCTGGTGTTTGGGAATATCTGGGGTGG	988
OY	301	AATGTTCAAGCCCTGTTGTTTGAAGAACTCATGTTGCTGAGTATCTCACTTCAAGAA	360
Db	989	AATGTGCACATGCTGTGTTGTTGAATGAGCTGTTCTGCTGAGTATCTGCTTTCAAGAG	1048
OY	361	GAGCTTGTGATGGAAGTTCAATGGAACCTTGTGTTTGAATTTGAATTTGAGCCCTTC	420
Db	1049	GAGCTTGTGAGGGAAGTTCTATATGGCAACTTTGTGCTTGAGTTGCACTTTGAACCGTTC	1108
OY	421	AAGCATCAATCCCCCGGCCCACTCTTTCAAAATCCATTTGTAATAGTGTGAAGTTCTCTA	480
Db	1109	AATGATCTTCTTCCTGGCCCACTGTGAACAATGTCATTTGAAATGGGTGAGTTCTTC	1168
OY	481	AATGCTCACTTTCGGCAAAATTTGTTGCATGACAAGAGAGCATCACCCCTTGTGCTGAA	540
Db	1169	AAACGCCACCTTTTGGGCCAAGCTCTTCATGACAAAGAGAGATGCACACCTGTTGAA	1228
OY	541	TTCTCTCAAGTCCATTGTTCACAAAGGCAAGAACATGATGTTGAATGACAAATTCAGAAC	600
Db	1229	TTCTCTCAGGCTTCAAGATTATTAAGGGAAGAACATGATGTTGAATGAACAAAGTTCAAGAC	1288
OY	601	TTGAATGCTCTTCAACATGTTTGTGAGAAAGCAGAGAGATCTGTGATCCCTACCTGCT	660
Db	1289	CTGGATTTCTTCCACACATGTTTGTGAAAAGCAGAAAGATCTGACTTCAGTTGCTCTCT	1348
OY	661	GAGACACCATGTCCCGAATTCGAAACACCGGTTCCAGAAATCGTTTGGAAAAGAGTTGG	720
Db	1349	GAAACACCCCTACTCGAATTTGAAAAACAATTCGGGAAATGGTTTGAAGAGGGGGTGG	1408
OY	721	GGTGACACCGCAAGACGGGTGTCGAGATGATCCACTCTTTTGGATCTTCTTGAAGCA	780
Db	1409	GGTGCATACGGCCGAGCGGTCTCTCGAATGATCCAGCTCTCTTGGACCTTCTTGAAGCA	1468
OY	781	ACTGATCTCTTGCAACCTTGAGAAGTTCTTGGGAGAAATCCCATGGTGTCAATTTTGG	840
Db	1469	CCCGACCTTTGCAACCTTCGAGACATTTCTTGGAAAGATTCCTATGTCTTCAATGTGTT	1528
OY	841	ATTCTCACTCCCAACGATATCTTGGCTCAAGACAATGTTTTGGGGTATCCCGACCGGT	900
Db	1529	ATCTCTTCTCCCAAGTATTACTTTTCCCAAGATATGCTTGGGGTACCTGACACTGGT	1588
OY	901	GGGCAAGTTGTTTAAATCTTGGATCAAGTTCGTGCTTGGAGAAAGAAATCTCAACCGC	960
Db	1589	GGACAGTGTGTTTAAATCTTGGATCAAGTTCGTGCTTGGAGAAAGAAATCTCAACCGC	1648
OY	961	ATTAAGCAACAGAGACTCAACATCACCCCTCGAATCTCATTAATACAGACTTCTTCTCT	1020
Db	1649	ATCAAGAAACAGAGCTTGATATCAACCCCTCGTATTTCTCATTAATCTGCTTCTCTCT	1708
OY	1021	GATGCTGTGGAACAACATGCGGCTCAACGACTTGAAGAAAGTAAAGCAACAGACACTCG	1080
Db	1709	GATGCAGTAGGAACATACCTGTGGCCAAAGCTTGAAGAGGATATAGATACGAATATGTGT	1768
OY	1081	GATATTTCTTCAAGTACCTTCAAGACAAAGAAAGGAATTTGTTCGAAATTTGATCTCAAGA	1140
Db	1769	GACATTTCTCGAGTTCCTTTCAAGAACGAAAGGGAATTTGTTGCGAAATTTGATCTCAAGA	1828

QY	1141	TTTGAAAAAGCTGCGCCATCTACTGGAAACCTACACAGAGATGTGGCTCATGAAATCTCC	1200
Db	1829	TTTCG---AGGCTGCGCATACCTAGAACTTACCTAGAGATGTGGCTTGGAACTTGGCC	1885
QY	1201	AAAGAGTTGACCGGACCGCCAGATCTGATCATTCGAAACNACAGCGACCGCAATATGCTC	1260
Db	1886	AAGAGTTGCAAGCCAAAGCCAGATCTGATCGTTGGAACTACAGTGAATGGAAACATTGTT	1945
QY	1261	GCTTCCTTGTCCGACATAAATTAGGTGTCACACAGTGCACCATCGCCCATGCTTTGGAG	1320
Db	1946	GCCTCTTTGTTAGACATTAATTAGAGTAACTCAGTGTACCAATTGGCTCATGCTCTAGAA	2005
QY	1321	AAGCAAAATATCCAGATTCAGATATCTATTGGAAAGCTTGAAGACAAATACCATTTC	1380
Db	2006	AAGACCAAGTACCCTGAGCTGCATTATTACTGGAAAAAATTGAAAGAAATATCACCTTC	2065
QY	1381	TCTTGSCAATTTACGCGATCTTTTGGCAATGAACCATACAGATTTCATCATACAGAT	1440
Db	2066	TCATGCCAATTACTGTGCTGATCTTTTGGCAATGAACCAACAGACTTTATCATCACAGC	2125
QY	1441	ACTTTCAGAAATTTGAGAGAAAGCAAGACACTGTGGTCAATACGAGACCACTGCT	1500
Db	2126	ACCTTCCAAAGATTTGCTGGAAGCAAGACACTGTTGGAACAGTATGAGAGTCACACTGCC	2185
QY	1501	TTCACTCTTCTGCTCTACCGTGTGTTACATGATTCAGTGTGTTGATCCCAAATTC	1560
Db	2186	TTCAACCTCTCCAGGACCTTACCGGTGTGTTCAACGGTATTCATCCCTTGTATCCAAAGTTC	2245
QY	1561	AACTTTGTTCCCTCGGTGCTGATATAGAGATATACCTCCCTTACCGCAAGAAACGCG	1620
Db	2246	AAACTTCGCTCTCCCGGTCGGAATAGACATATATCTCCATACCTGAAACTGAGGCT	2305
QY	1621	AGTTTGAAGACTTTCACATCCAGATGAAAGACCTTCTTTACACCAAGTTGAGAAATGAA	1680
Db	2306	AGGTTAACAGAGTTCCACCCCGACATTTAGAGGCTTTTACACACTCAGTGGAAATGAA	2365
QY	1681	GAACACTTATGTGTCTCATATGACCGCAACAGCCAAATTCGTTCACATGSCCAAGCTT	1740
Db	2366	GAACACATATGTATTTAGAAAGACCGCAACAAAGCCGATCATCTTCACCATGSCCAAGCTT	2425
QY	1741	GATGTGTCAAGAACTTAAACCGGACTGTGTGAGTGTGCGGCAAGAACCCAAATTCGCT	1800
Db	2426	GACCGTGTGAAGAACATCACGGGACTGTGTGAGTGTATGGCAAGAAATGCGCGCTCCGC	2485
QY	1801	GAGTTGGCTAACTCCGTAGTTGTATGAGTGTGATAGCGAAAGAAAGATCTAAAGATTGGAA	1860
Db	2486	GAGTTGGTAAACCTCGTGTGTGTGGCCGGAGACAGAGAAAGTGTCCAAAGACTTGGAA	2545
QY	1861	GAGAAAGCTGAATGAAGAAATGTTTGAAGTATGACGTATGACATGAACTTGAACCGGCCAA	1920
Db	2546	GAGAAAGCCCGAGATGAAGAAAGATGTATGAGCTTACATGAGACCTTCAAGTTGAAACGGCCAA	2605
QY	1921	TTTCAGATGAGATATCATCTCAATATGAACAGAAATCCGAATGTGAACTTTACCGATACATT	1980
Db	2606	TTTCAGATGAGATCTCTCTCAGATGAATACCGTGTGAGAAACGAGAGCTCTACCGTGTATC	2665
QY	1981	TGCACACAGAAAGGTGCTTTGTATCAGACTGTCATGTATGAAGGCTTTGGATTGACAGTT	2040
Db	2666	TGTATCACAAAGGGTGTCTTTGTGTACAGCTGTGCACTGTATATGAGGCTTTGGGTTGACGTG	2725
QY	2041	GTGAGGCAATGACTTSCGTTTGGCAACATTCGCAACTGTAAACGCTGTACAGCCGAG	2100
Db	2726	GTTGAGGCGCATGACTTGTGGGTATCCAAACTTTGCCACATGCAATGTGTGCTGTGAG	2785
QY	2101	ATTATTTGTCCATGGGAATCTGTGTTTCAACATTTGATCTTTACATGTGTATCAAGCTGCT	2160
Db	2786	ATCATTTGTGACATGGAAATCTGTGTTTCAACATTTGACCTTTACCAATGTGTACCGTGTGCT	2845
QY	2161	GACATPACTCGTCGATTTCTTTGAAAGGTATGAAGAAATCAATCTCACCTGGAGATAGAGT	2220
Db	2846	GAGATCTCTGTGTGATCTTTTGAAGAAAGCAAGGCTGATCCATCTTCACTGGGACAAATATC	2905
QY	2221	TCCCAAGAGGCTTGAACGATATGAGAGAAAGTATACATGGAAGATTTTACTCGAGAGA	2280

Db	2906	TCCACGGGTGACCTCAGACCGTATTCATGAGAAAGTACACATGCGCAATTACTCTGACAGG	29655
Qy	2281	CTATTCACCCCTGACAGAGAGTATGATTCGGAAGCATGTTTCCAACTTGAACGCCGT	23440
Db	2966	CTCTTGACACTCACCTGCTGTGTATATGGCTTCTCGAAGCATGTGACCAACTTGGAAAGCCGT	30255
Qy	2341	GAGAGTCGTCCCTTACCTTGAGATGTTTATGCTCTTAAATGACCCGTAGCTGCTGAATCA	24000
Db	3026	GAGAGCAAACTTACCTTGAGATGTTCTATGCTCTCAGATGACCGCAATTGGCTGAGCTT	30855
Qy	2401	GTTCCATTGCGCAGAGAGAGTAATTGAACTGTAAATACA	2441
Db	3086	GTGCCCTCTGCTATTGGAAGAGATAAATATATGTTTGAAGAGA	3126

RESULT 8
US-10-425-114-9594

APPLICANT: Jiu, Jingdong
 APPLICANT: Zhou, Yihua
 APPLICANT: Kovalic, David K.
 APPLICANT: Screen, Steven E
 APPLICANT: Tabaska, Jack E
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

OTHER INFORMATION: Clone ID: 700852649_FLI
US-10-425-114-9594

Query Match	59.7%	Score	1566.4	DB	13	Length	2772	
Best Local Similarity	77.9%	Pred.	No. 0					
Matches 1901	Conservative	0	Mismatches	537	Indels	3	Gaps	1.

QY	1	TTGGGTGAGCGGCTCTCACTCGCGGTCACAGCTCCGGAGGGCTTTGGATGAGACCCCTT	60
Db	53	ATGGGAATCAACCCCTTTGACACCTCTCACTCTTCGGAGAGGTTTGATGAACTCTC	112
QY	61	CTTGCTCAAGAGACGAGATTTTGGCCTTGCTCAAGANTCGAGGGCAAGGAAAAGA	120
Db	113	ACTGGTCACAGGAATGAAATTTTGGCCCTTTTGTCGAAGGCTTGAAOCCAAAGGCCAAGGA	172
QY	121	ATTCTGCAACACCATCAAAATTAATTCTAGAGTTTGAAGCTATCCCTGAAGAGAACAGAAAG	180
Db	173	ATCTGCAACACACACAGGTGGTGGAGAGTTTGAAGAAATCCCTGAAGAGAACAGAAAG	232
QY	181	AAAGTCGGCTAATGGTGCATTTTGTGAAGATTAAGAGCTAAGTACAGAAAGCATCGCTTG	240
Db	233	AAAGCTCCAAAGGTGGTCTTTGGAGAAAGTTTGAGATCTACACAGAAAGCATATGGCTG	292
QY	241	CCTCCATGGGTGCACCTTGCTGTTCGTCGAAGGCCGTGTGTTGGAGATCATTAAGTG	300
Db	293	CCACCATTTTGTGGCTCTGGCTGTTCGACCAAGGCCGTGTGTTTGGAGATATCTGGCGGTG	352
QY	301	AATGTTCAAGCCCTTGTGTTGTAAGAACTCATGTTGCTGATATCTCCACTTCAGAGAA	360
Db	353	AATGTCACATCTGTTGTGTAATGATGCTCTCTGCTGAGATATCTGCGTTTCAAGAG	412
QY	361	GAGCTGTGTAAGGAAGTTCAAAATGGAACCTTGTTTGGAAATTTGATTTGAGCCCTTC	420
Db	413	GAGCTGTGTGAGGAAGTTCTATATGCAACTTGTGTCTGTAGTTGACCTTTGAACCGTTT	472

QY	421	AACATCATTTCCCCCGCCCACTCTTTCAAAAATCAATGGGTATGGTGTGAATCTTCTTA	480
Db	473	AATGATATCTTCTCCGCGCCAACTGTGAACAAGTCAATTGGAAATGGCGTGTGAATCTCTC	532
QY	481	AATGTCAACCTTTGGGCAAAATTTGTTCCATGTACAAGAGAGATGACACCTTTTGTCTGGAA	540
Db	533	AACGGCCACCTTTGGGCCAAGCTCTTCATGTACAAAGAGAGATGACAGCCACTGCTTGGAA	592
QY	541	TTCTCTCAGAGTCCATTGTCAACAAGGCGCAAGAACATGATGTTGAATGACAGAAATTCAGAAC	600
Db	593	TTCTCTCAGGCTTCCAGATTATTAAGGAAAGCCATGATGTTGAATGTCAAAAGTTCAAAGC	652
QY	601	TTGAATGCTCTTCAACATGTTTTTGAAGAAAGACAGAGAGATATCTGTGATCCCTACCTCT	660
Db	653	CTGAATTTCTTCCAGCAAGTGTTTTGAAGAAAGCAAGAGATGTCTGACTTCAATGTCTCTCT	712
QY	661	GAGACACCATATGCGCGCAATTCGAACACCGGTTCCAGAGAAATCGGTTTGAAGAAGGTTGG	720
Db	713	GAATCACCCCTACTCGAATTTGAAATAAATAATCCGGAAATGGTTTGGAGAGGGGGGTGG	772
QY	721	GGTACACCGCAGAACGCGTGTCTGAAGATATCCAATCTCTTTTGAATCTTCTTTAGGCA	780
Db	773	GGTACATATCGCGCAGCGTGTCTCGAATATATCAAGTCTCTCTTGAACCTTCTTGAAGCA	832
QY	781	ACTGATCTTGGACACCTTTGAGAAATTTCTTGGGGAATCCCAATGATGTCAATGTTGTG	840
Db	833	CCCAACCTTTGACACCTCTGAGACATTTCTTGGAAAGATTCCTATGTCTTCAATGTTGTT	892
QY	841	ATTCTCACTCCCAACGGATATCTTGTCTCAAGACATGTTTGGGGGTATCCCGACACCGGT	900
Db	893	ATCTTTCTTCCCAATGGTTACTTTTGGCCAAATATGTCTTGGGGTATCCCGACATCTGTT	952
QY	901	GGGACGTTGTTTATCATCTTGGATCAAGTCCGAGCTTTGAGAAATGATATGCTCTCTCGT	960
Db	953	GGAAGGTTGTTTATCATCTTGGATCAAGTCTGCTCTTGGAGAAATGATATGATGCTCAACGCG	1012
QY	961	ATTAAGCAACAAGACATCAACATCAACCCCTCGAATCTCTCATTTACTATAGACTTCTTCCCT	1020
Db	1013	ATCAAGAAACAAGGCTGTGATATCAACCCCTCGATTTCTCATTTACTCTGCTTCTCTCT	1072
QY	1021	GATGCTGTGGAAACAATGTCGGTCAACGACTTGAGAAAGTATCGGAACAGACACTCG	1080
Db	1073	GATGACGTAGGAATCTACTGTGGCCAAACGTGTAGAGGGGTATGATCTGAATATTTGT	1132
QY	1081	GATATCTTTCGAGTACCTCTTCAGAACAAAGGAAATTTGTTCGAAATGATCTCAGA	1140
Db	1133	GACATTTCTCGAGTTCTCTTTCAGAACCGAAAGGAATTTGTTCGAAATGATCTCAGA	1192
QY	1141	TTTGGAAAAAGTCTGGCCATTACTTGGAAACCTTACACAGAGATGTTGCTCATGAATCTCC	1200
Db	1193	TTTCG---AAGTCTGGCCATATCTTGAAGACTTACCTAGAGATGTTGCTTGAACCTTGGCC	1249
QY	1201	AAAGAGTTTGCACGGCACGGCCAGATCTGATATCATGGAAACNACAGCGACGGCAATATGCTC	1260
Db	1250	AAGGAATTTGCACAGCCAAAGCCAGATCTGATGTTTGGAACTTACAGTGTATGGAAACATTTGTT	1309
QY	1261	GCCTCTTGTCTCGCACATAATTATTTAGTGTCAACAGTGCACCATCGCCCATGCTTTTGGAG	1320
Db	1310	GCCTCTGTTTGTAGACATTAATTATTTAGAGTAACTCAAGTATACANTTGTCAATGCTCTTGA	1369
QY	1321	AAGACAAATATATTCAGATTTCAATATCTATTTGGAAAGAGCTTGAAGCAAAATACATTTTC	1380
Db	1370	AAGACAAAGTACCTGTAGTCTGACATTTTACTGGAAGAAATTTGAAAGAAATATACATTC	1429
QY	1381	TCTTGGCAATTTAAGCGATATCTTTTGGCATATGAACCATACAGTTTCATCATACACAGT	1440
Db	1430	TCAATGCAATTTTACTGCTGATCTTTTGGCATATGAACCAACAGACTTTATCATACAGC	1489
QY	1441	ACTTTCAGAGAAATTTGAGAGAGACAGACACTGTTGTCAATATCGAGACCAACTGCT	1500
Db	1490	ACCTTTCAGAGAGATTTGCTGAGAGACAGACACTGTTGTGAACATATGAGATGACACACTGCC	1549
QY	1501	TTCACTCTTCTGTGCTCTATCCGTGTTGTATCATGTATGATATGATATCCAAATTC	1560

781 ACTGATCTGACCCCTTGAGAAAGTTCCCTGGAGAAATCCCAGTGTGTAATGTTG 840
1013 CCCGACCCGTCACCTCTCGAGAAAGTTCTTGATAGAGTTCCCAAGTCTTCAAGCTG 1072
841 ATTCTCATCCCAAGGAAATCTTCTGCTCAAGCAATGTTTTGGGTATCCGCAACCGGT 900
1073 ATCATGTCCTCCCAAGGAAATCTTGTCTCAGGACAGAGTCTTGTGTTATCCGGAATCCGGT 1132
901 GGGCAGGTTGTTTACATCTTGATCAAGTCCGAGCTTTGGAGAAATGAGATGCTCTCCGT 960
1133 GGGCAGGTTGTTTACATCTTGATCAAGTCTTGTGTTATCCGGAATCCGGT 1192
961 ATTAAGCAACAGGAACTGAACTACCCCTGAAATCTCATTAATTAATGATCTTCT 1020
1193 ATTAAGCAACAGGAACTGAACTACCCCTGAAATCTCATTAATTAATGATCTTCT 1252
1021 GATGCTGTGGAACAACTGCGGTCAAGCACTTGAGAAAGTATACGAAACAGAGCACTG 1080
1253 GACGCGTTGGAAACCACTGTGGCCAGAGGCTTGAGAAAGTTTGGGACCGAGTACTC 1312
1081 GATATTTCTTCAAGTACCTTCAAGCAAGAAAGGAAATGTTGAAATGATCTCAAGA 1140
1313 CACATTTCTTGGGTCCTTCAAGAAATGAGAGGAGTGTCCGCAAGTGGATTTCCCGG 1372
1141 TTTGAAAAGTCTGGCAATCTTGAAACCTACAGAGAGATGTTGCTCATGAAATCTCC 1200
1373 TTCC--AGGTGTGGCCCTTATTTGAAAGATACCTGAGAGATGTCGACGAACTTGTCT 1429
1201 AAAGAGTTGACGCGACCCAGATCTGATCATCGGAAACNACAGCGCAATATCTC 1260
1430 GAGAGTTGACGCGGAGGAGCTGATCTGATCATCGGAAACNACAGAGATGAGAACTGTT 1489
1261 GCTCTCTTGTGCAATTAATTAAGTGTCAACAGTSCAACATCGCCCATGCTTGGAG 1320
1490 GCTTCTCTGTAGCACTAATTAAGTGTACACAGTGTACAAATGACCCTCCAG 1549
1321 AAGCAAAATATCCGATTCATATCTATGAGAAAGCTTGAAGAAATATCAATTC 1380
1550 AAGCAAAATATCCGATTCATATCTATGAGAAAGCTTGAAGAAATATCAATTC 1609
1381 TCTTGCCAAATTAAGTATCTTTTGCATGAACATACAGATTCATCATCAGT 1440
1610 TCTTGCCAAATTAAGTATCTTTTGCATGAACATACAGATTCATCATCAGT 1669
1441 ACTTTCAGAAATTAAGTATCTTTTGCATGAACATACAGATTCATCATCAGT 1500
1670 ACTTTCAGAAATTAAGTATCTTTTGCATGAACATACAGATTCATCATCAGT 1729
1501 TTTCACTCTTCTGCTGCTCTACCGGTGTGTAATGATGATGTTGATCCCAATTC 1560
1730 TTTCACTCTTCTGCTGCTCTACCGGTGTGTAATGATGATGTTGATCCCAATTC 1789
1561 AACATTTGTTCCCGTGTGCTGATATGAGATATCTTCCCTTACACGGAAGAGGCG 1620
1790 AACATTTGTTCCCGTGTGCTGATATGAGATATCTTCCCTTACACGGAAGAGGCG 1849
1621 AAGTTGAAGATTTTCAATCTGAGATGAGAGACTTCTTTTACACCAAAAGTTGAGATGA 1680
1850 CGGTGGAATCTTCCACCTGATGAGAGAACTCTTTCAGGAGATGAGAGAAAG 1909
1681 GAGCACTTATGTGTCTCAATGACGCAACCAAGCAATTCCTTGTTCACAAATGCAAGGCTT 1740
1910 GAGCACTTATGTGTCTCAATGAGAGAGGCTTATTTTACCAATGCAAGGCTG 1969
1741 GATCGTGTCAAGAACTTAACCGGACTCGTCAAGTGTGCGGAGAAACCAAGTGTGCT 1800
1970 GACCTGTGTCAAGAACTTAACCGGACTGTGTGATGTGTATGAGAGAACTTCAAGTGTGAG 2029
1801 GAGTTGCTTAACCTGTAGTGTGATGTGTGATGAGAGAAATCTTAAGATTTGAA 1860
2030 GAACTCGCAACTGTGTGTGTGAGAGTGAAGAGAGATTTCAAGAGACTTGAA 2089

1861 GAGAGGCTGAAATGAGAAATGTTGAGCTGATGACAGTACAACTTGACGCCAA 1920
2090 GAGCAGCTGATGAGAAATGTTGAGCTGATGACAGTACAACTTGACGCCAA 2149
1921 TTCAATGATATCATCTCAATGAGAGATCCGAAATGTTGAATTTACCATTAAT 1980
2150 TTCAATGATATCTTCCCAAGATGAGAGCTGATGAGAGATGAGAGCTTACCATC 2209
1981 TGGACACGAAAGGCTTGTGACAGCTGATGATGAGAGCTTGTGATGAGCTT 2040
2210 TGTACACGAAAGGCTTGTGACAGCTGATGATGAGAGCTTGTGATGAGCTT 2269
2041 GTGAGGCAATGATCTTGTGAGCTTGTGACAGCTTGTGACAGCTTGTGACAG 2100
2270 GTGAGGCAATGATCTTGTGAGCTTGTGACAGCTTGTGACAGCTTGTGACAG 2329
2101 ATTAATGCTGAGGAAATCTGTTTCAATGATCTTGTGACAGCTTGTGACAG 2160
2330 ATCAATGCTGAGGAAATCTGTTTCAATGATCTTGTGACAGCTTGTGACAG 2389
2161 GACATCTGCTGAGGAAATCTGTTTCAATGATCTTGTGACAGCTTGTGACAG 2220
2390 GAGCTTCTTGTGAGCTTGTGACAGGCTTGTGACAGCTTGTGACAGCTTGTG 2449
2221 TCCCAAGGAGCTTGAAGCAATGAGAGAGTATGATGAGAGATTTTCTCGAGAG 2280
2450 TCAAGGCTGCTGAGGAGTATGATGATGAGAGATTTTCTCGAGAG 2509
2281 CTATGACCTTGAAGAGTATGATGATGAGAGATTTTCTCGAGAG 2340
2510 CTGTTGAACCTGACCTGCTGTATGCTTGTGAGAGATGAGAGATTTTCTCGAG 2569
2341 GAGAGTGTGCTTACCTTGAAGATTTTATGCTTGTGAGAGATTTTCTCGAG 2400
2570 GAGAGTGTGCTTACCTTGAAGATTTTATGCTTGTGAGAGATTTTCTCGAG 2629
2401 GTTCATTTGAGAGAGATTA 2422
2630 GTTCTCGGCTGTGAGATTA 2651

RESULT 10
US-10-393-840-15
; Sequence 15, Application US/10393840
; Publication No. US20030229922A1
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1012c3
; CURRENT APPLICATION NUMBER: US/10/393,840
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 09/636,800
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/170,862
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: US 60/148,426
; PRIOR FILING DATE: 1999-08-11
; PRIOR APPLICATION NUMBER: PCR NZ/99/00169
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 956
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 2913
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-393-840-15

Query Match 57.1%; Score 1497.8; DB 16; Length 2913;
Best Local Similarity 76.4%; Pred. No. 0;
Matches 1851; Conservative 0; Mismatches 568; Indels 3; Gaps 1;
1 ATGCTGAGCGTGTCTCATCTCGGTCACAGTCTCGGTGAGCTTGTGATGAGACCTT 60


```
Db 242 ATGGGTGATTCAGTGTACTCGAAGCCAGACGCTTCCGAGCGTTTGAGAGACCTCTC 301
Qy 61 CTTCCTCACAGAAACGAGATTTTGGCTTCCTCAAGATCGAGGCAAAAGAAAGA 120
Db 302 TCTGTCAACCGCAACGATATTTGGCTTCCTTCAAGGTTGAAGCCAAAGGCAAGGC 361
Qy 121 ATTCGCAACCAATCAATATTTAGAGTTGAAGCTATCTCGAAGAAACAGAAAG 180
Db 362 ATCTTGACAGCCCAACAAATTTTCTGAGTTTGAGGCCATCTCGAAGAAAGCAGAC 421
Qy 181 AAGCTCGTAAATGATGCAATTTTGAAGTAAAGCTATGACGGAAGCAATCGTGTG 240
Db 422 AAGCTCTTGAGTGGGCTTTGGTGAAGTCTCAAAATCCTCAGAGAAAGCAATGTGTG 481
Qy 241 CCTCATGAGGTTGCACTTGCTGTCACAAAGCCTGTGTTGGAGTATAGATG 300
Db 482 CCTCATGAGGTTGCTCTTGCTGTCACAAAGCCGGGCGTGTGGAGACATCCGTGTG 541
Qy 301 AATGTCAAGCCCTGTTGTTGAGGAACTCACTGTTGATGATCTCCACTTCAAGAA 360
Db 542 AAGCTCAAGCCCTGTTGTTGAGGAAATGGAAGTTGCTGAGTATCTGCACTCAAGAA 601
Qy 361 GAGCTTGTGATGATGATTCAAATGGAACCTTGTGTTGGAATTTGATTTGAGCCCTTC 420
Db 602 GAGCTTGTGATGATGATTCAAATGGAACCTTGTGTTGGAATTTGATTTGAGCCATTC 661
Qy 421 AACTCATATTTCCCGCCCAACTCTTCAAAATCATTTGATATGATGATGATTCCTA 480
Db 662 ACTGCTCTTTTCCCGCCCAACTCTTCAAAATCATTTGATATGATGATGATGATTC 721
Qy 481 AATGTCAACCTTTCGCAAAATTTGTCATGACAAAGAGATGACACCTTTGCTGAA 540
Db 722 AATGTCAACCTTTCGCAAAATTTGTCATGACAAAGAGATGACACCTTTGCTGAA 781
Qy 541 TTCTCAAGTTCATGTCACAAAGGCAAGAAATGATGTAATGACAAATTCAGAAC 600
Db 782 TTCTCAAGTTCATGTCACAAAGGCAAGAAATGATGTAATGACAAATTCAGAAC 841
Qy 601 TTGAATGCTCTTCAACATGTTTGAAGAAAGAGAGATATCTTGTATACCTTCTCT 660
Db 842 GTGTTCTCTCTCAACATGTTTGAAGAAAGAGAGATATCTTGTATACCTTCTCT 901
Qy 661 GAGACACCATGTGCGGAATTCGAACACCGGTTCCAGAAATCGGTTTGAAGAGTTGG 720
Db 902 GAGACACCATGTGCGGAATTCGAACACCGGTTCCAGAAATCGGTTTGAAGAGTTGG 961
Qy 721 GGTGACACCGGAGAAACGGGTGCTGAGATGATCCAACTCTTTGGAATCTTTGAAGCA 780
Db 962 GGTGACACCGGAGAAACGGGTGCTGAGATGATCCAACTCTTTGGAATCTTTGAAGCA 1021
Qy 781 ACTGATCTTGCACACCTTGAAGATTCCTTGGAGAAATCCCATGATGATGATGATG 840
Db 1022 CCCGACCCGTGACCTCGAAGATTCCTTGAAGATTCCTTGAAGATTCCTTGAAGAT 1081
Qy 841 ATCTCACTCCCAACGATACTTGTCTCAAGAAATGTTTGGGATATCCGACACCGGT 900
Db 1082 ATCTCACTCCCAACGATACTTGTCTCAAGAAATGTTTGGGATATCCGACACCGGT 960
Qy 901 GGGCAGGTTGTTAATCTTGAATCAAGTCCGAGTTTGAAGATGATGATGATGATG 960
Db 1142 GGGCAGGTTGTTAATCTTGAATCAAGTCCGAGTTTGAAGATGATGATGATGATG 1201
Qy 961 ATAAAGCAACAGATCAATCAACCTCGAATCTCATTAATATAGACTTCTTCT 1020
Db 1202 ATTAAGCAACAGATCAATCAACCTCGAATCTCATTAATATAGACTTCTTCT 1261
Qy 1021 GATGCTGTGGAACAAATGCGGTCAACGATTCGAAAGTATACGAAACAGACATCG 1080
Db 1262 GACGCGTGTGGAACAAATGCGGTCAACGATTCGAAAGTATACGAAACAGACATCG 1321
Qy 1081 GATATTTCTTGAAGTACCTTGAAGCAAGAAAGGAATTTGTCAAATGATCTCAAGA 1140
Db 1322 CACATTTCTGCGCTCCCTTCAAGAAATGAGAAAGGATCTGTCGCAATGATTTCCCG 1381
Qy 1141 TTGAAAAAAGCTGAGCAATCTTGAAGAACTTACACAGAGATGTTGCTCATGAATCTCC 1200
Db 1382 TTGCG--AGTGTGCGCTATTTGAAAGATACCTGAGGATGTGCGAGGAATCTGCT 1438
Qy 1201 AAAGATTCGACGCGACCGCATCTGATCATCGGAAACNACAGCGCATATGTCTC 1260
Db 1439 GGAAGTTGCAAGGCAAGCCGATCTGATCATCGGAAACNACAGCGCATATGTCTC 1498
Qy 1261 GCGTCTTGTGCGCAATTAATTAGTGTCAACAGTGCACATCGCCCATGCTTTGGAG 1320
Db 1499 GCTTCTTGTGAGCAATTAATTAGTGTTCACAGTGTACAAATAGCCCATGCTTGGAG 1558
Qy 1321 AAGCAAAATATCCAGATTCAGATATCTATTTGAAGAAAGCTTGAAGCAATATCATTTTC 1380
Db 1559 AAGCAAAATATCCAGATTCAGATATCTATTTGAAGAAATTTGAAGAAATATCATTTTC 1618
Qy 1381 TCTTGCCAAATTTACAGCTGATCTTTTGAAGTAAACCATATCATCATCATCAT 1440
Db 1619 TCTTGCCAAATTTACAGCTGATCTTTTGAAGTAAACCATATCATCATCATCATCAT 1678
Qy 1441 ACTTTCAGGAAATTTGAGGAGCAAGGACATGTTGATCAATAGAGGCAACATGCT 1500
Db 1679 ACTTTCAGGAAATTTGAGGAGCAAGGACATGTTGATCAATAGAGGCAACATGCT 1738
Qy 1501 TTCACTCTTCTGCTCTCAACCTGTTGTAATGATATGATATGATGATGATGATGATG 1560
Db 1739 TTCACTCTTCTGCTCTCAACCTGTTGTAATGATATGATATGATGATGATGATGATG 1798
Qy 1561 AACATTTGTTTCCCTGCTGATATGAGATATATCTTCTTCAACCGAAGAAAGCGG 1620
Db 1799 AACATTTGTTTCCCTGCTGATATGAGATATATCTTCTTCAACCGAAGAAAGCGG 1858
Qy 1621 AGTTGAAGATTTCCATCTGATGATGAGATGAGATGATGATGATGATGATGATG 1680
Db 1859 CGTTGAAGATTTCCATCTGATGATGAGATGAGATGATGATGATGATGATGATG 1918
Qy 1681 GAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
Db 1919 GAACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1978
Qy 1741 GATGTTGCAAGAACTTAAACCGACCTGTCAGATGATGATGATGATGATGATGATG 1800
Db 1979 GATGTTGCAAGAACTTAAACCGACCTGTCAGATGATGATGATGATGATGATGATG 2038
Qy 1801 GAGTTGCTAACCTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
Db 2039 GAGTTGCTAACCTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 2098
Qy 1861 GAGAAAGCTGAATGAAGAAATGTTTGAAGTGAATGCAATGCAATGCAATGCAAT 1920
Db 2099 GAGAAAGCTGAATGAAGAAATGTTTGAAGTGAATGCAATGCAATGCAATGCAAT 2158
Qy 1921 TTCAATGATATATCTCAATGAACAGATTCGAAATGTTGAATCTTACCTATCATTT 1980
Db 2159 TTCAATGATATATCTCAATGAACAGATTCGAAATGTTGAATCTTACCTATCATTT 2218
Qy 1981 TGACACAGAAAGGTCCTTGTGACAGCTGATGATGATGATGATGATGATGATGATG 2040
Db 2219 TGACACAGAAAGGTCCTTGTGACAGCTGATGATGATGATGATGATGATGATGATG 2278
Qy 2041 GTGAGGCAATGATCTTGGGTTTGCACAACTTGCACACTGTAAACGTTGACAGCGAG 2100
Db 2279 GTGAGGCAATGATCTTGGGTTTGCACAACTTGCACACTGTAAACGTTGACAGCGAG 2338
Qy 2101 ATTAATGATGAGAAATCTGTTTCAACATTTATCTTATCTTATCTTATCTTATCTT 2160
Db 2339 ATTAATGATGAGAAATCTGTTTCAACATTTATCTTATCTTATCTTATCTTATCTT 2398
Qy 2161 GACATATCTGTCATCTTCTTGAAGATGTAAGAAAGTATCATCTCATGAGATTAAGATC 2220
Db 2399 GACATATCTGTCATCTTCTTGAAGATGTAAGAAAGTATCATCTCATGAGATTAAGATC 2458
```


2221 TCCCAAGAGGCTTGAAGCAATAGAGAGATATACATGAGATTACTCGAGAGA 2280
2459 TCAGAGGCTGCATCAGAGATTAAGAGAGATATACATGAGAAATATATCTGAGAG 2518
2281 CTATTGACCTGACAGAGATGATGATTTCTGGAAGCATGTTTCAACCTTGAGCCGT 2340
2519 CTGTTGAACCTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2578
2341 GAGAGTCTGCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
2579 GAGAGTCTGCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2638
2401 GTTCATTTGGCAGAGAGATAA 2422
2639 GTTCTCCGCTGTCAGATGATA 2660

RESULT 11
US-10-137-036-57
Sequence 57, Application US/10137036
Publication No. US20030101478A1
GENERAL INFORMATION:
APPLICANT: Perera, Ranjan
APPLICANT: Rice, Stephen
APPLICANT: Esigle, Clare
APPLICANT: Leaham, Annette
APPLICANT: Wood, Marion
APPLICANT: Visser, Elizabeth
TITLE OF INVENTION: Compositions and Methods for the
FILE REFERENCE: 11000.10364
CURRENT APPLICATION NUMBER: US/10137, 036
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: PCT/NZ 01/00115
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/724,624
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/598,401
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: PCT/NZ00/00018
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 60/146,591
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: U.S. No. US20030101478A1 09/276,599
NUMBER OF SEQ ID NOS: 143
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 57
LENGTH: 3103
TYPE: DNA
ORGANISM: Eucalyptus grandis
US-10-137-036-57

Query Match 57.1% Score 1497.8; DB 15; Length 3103;
Best Local Similarity 76.4%; Pred. No. 0;
Matches 1851; Conservative 0; Mismatches 568; Indels 3; Gaps 1;

QY 1 ATGCGTGAAGCTGCTCTCACTGCGTCGACAGTCTCCGTGAGCGTTGGATGAGACCTT 60
DB 447 ATGCGTGAAGCTGCTCTCACTGCGTCGACAGTCTCCGTGAGCGTTGGATGAGACCTT 506
QY 61 CTGCTGACAGAGAGAGATTTTGGCTTGGCTCTCAAGATGAGAGGCAAGAAAGGA 120
DB 507 TCTGCTACCGGAGAGATTTTGGCTTGGCTCTCAAGATGAGAGGCAAGAAAGGA 566
QY 121 ATTGCAACACCATCAATATTCTAGAGTTGAAGCTATCCCTGAGAGAGAGAAAG 180
DB 567 ATCTGCAAGCGGACCAAGATTTTGGCTTGGCTCTCAAGATGAGAGGCAAGAGCA 626
QY 181 AAGCTGCTAATGTCATTTTGAAGTGAAGGCTAGTCAGAGAGAGAGATGCTGTTG 240
DB 627 AAGCTTCTTGAAGGCGCTTGGTGAAGTCTCAATCACTCAGAGAGAGAGATGCTG 686

241 CTTCATGGGTTGCACTTGTCTGTCACAGGCGTGTGTTGGAGATTAAGAGT 300
667 CTTCAATGGGTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 746
301 AATGTCACGCGCTTGTGTTGAGAACTCACTGTGCTGAGTATCTCACTTCAAGAA 360
747 AAGCTGATGAGGCTTGTCTTGAAGAAATGAGGTTGTGATATCTCACTTCAAGAA 806
361 GAGCTTGTGATGAGAGTTCAATGAAACTTTGTTGAAATGAGATTTTGAAGCTTTC 420
807 GAGCTTGTGATGAGAGTTCAATGAAACTTTGTTGAAATGAGATTTTGAAGCTTTC 866
421 AACTCATGATTTCCCGCGCGCACTCTTCAAAATTCATTTGGAATGAGTGAAGTTCTA 480
867 AAGCTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 926
481 AATGTCACCTTTGCGCAAAATTTGTCATGACAGAGAGATGACACCTTTGCTGGA 540
927 AATGTCACCTTTGCGCAAAATTTGTCATGACAGAGAGATGACACCTTTGCTGGA 986
541 TTCTCAGAGTTCATGTCACAGGCAAGACATGATGATGATGATGATGATGATGATGAT 600
987 TTCTCAGAGTTCATGTCACAGGCAAGACATGATGATGATGATGATGATGATGATGAT 1046
601 TTGAATGCTTTCAACATGTTTGAAGAAAGCAGAGAGATGATGATGATGATGATGATG 660
1047 GTTGTCTCTTCCCAACATGTTTGAAGAAAGCAGAGAGATGATGATGATGATGATGATG 1106
661 GAGACACCATGTCGCAATTTGGAACACCGTTCAGAGAAATCGTTTGAAGAGTTGG 720
1107 GAGACACCATGTCGCAATTTGGAACACCGTTCAGAGAAATCGTTTGAAGAGTTGG 1166
721 GGTGACACCGGCAAGACCGGCTGAGATGATGATGATGATGATGATGATGATGATGATG 780
1167 GGTGACACCGGCAAGACCGGCTGAGATGATGATGATGATGATGATGATGATGATGATG 1226
781 ACTGATCTTGCACCTTGAAGAGTCTTGGAGAAATCCCATGATGATGATGATGATGATG 840
1227 CCGGACCGGCTGCTGCAATTTGGAACACCGTTCAGAGAAATCGTTTGAAGAGTTGG 1286
841 ATTGTCACCTTCCCAACATGTTTGAAGAAATCGTTTGAAGAGTTGG 900
1287 ATTGTCACCTTCCCAACATGTTTGAAGAAATCGTTTGAAGAGTTGG 1346
901 GGGCAGGTTGTTACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
1347 GGGCAGGTTGTTACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1406
961 ATTAAGCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
1407 ATTAAGCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1466
1021 GATGCTGTCGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
1467 GATGCTGTCGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1526
1081 GATGCTGTCGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
1527 GATGCTGTCGAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1586
1141 TTTGAAAAAGTGTGCTTATCTTGAAGAACTTCAAGAGAGATGATGATGATGATGATGATG 1200
1587 TTTGAAAAAGTGTGCTTATCTTGAAGAACTTCAAGAGAGATGATGATGATGATGATGATG 1643
1201 AAGAGTTGACAGGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
1644 GAGAGTTGACAGGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1703
1261 GCTTCTTGTGTCGACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
1704 GCTTCTTGTGTCGACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1763

QY 1321 AAGACAAATATCCAGATTCAGATATCTATTGGAAAGAGCTTGAAGCAATACCATTTC 1380
DB 1764 AAGACGAAGTACCAGAGTACAGATATATCTGAAGAAATTTGAGAAAGATACCACTTC 1823
QY 1381 TCTTGCCAAATTAACGCTGATCTTTTGCATGAAACCATACAGATTTTCATACCCAGT 1440
DB 1824 TCTTGCCAGTTCACTGCTGATCTTCATGCCATGAAACCAACCGATTCATATATCAACG 1883
QY 1441 ACTTTCAGAAATTCAGAAAGCAAGCACTGTGTCAATAAGAGGCCACTGCT 1500
DB 1884 ACTTTCAGAAATTCGTAAGAGCAAGATACAGTGGGCGATAGAGTCAACATGAC 1943
QY 1501 TTCACTTCTCTGTCTCTACCGTGTGTATCATGTGTATCATGTGTGATCCAAATTC 1560
DB 1944 TTCACTTCTCTGTCTCTACCGAGTTGTTCACCGGATTCAGATCTTTCGACCCGAAAGTTC 2003
QY 1561 AACATTTTTCCTCGTGTGTATGAGATATCTTCCCTTACACCGAAAGAGAGGG 1620
DB 2004 AACATTTTTCACCGAGTGTGTATGAGATATCTTGTGTATCACCGAAAGAGAGGG 2063
QY 1621 AGTTGAAGCAATTCATCTGAGATCGAAGACCTTCTTACACCAAGTTGAGATGAA 1680
DB 2064 CGTTGAAGATCTTCCACCTGAGATCGAAGATCTTCCAGGATGTTGAGAACAG 2123
QY 1681 GAACACTTATGTGTCTCATATGACCGCAACAAAGCCATTTCTTCAATGCGAAGCTT 1740
DB 2124 GAACACTTGTGTGTGTGAAGAATGAAGAAGCCATTTATTTTTCACCATGCAAGGCTG 2183
QY 1741 GATGTGTCAAGAACTTAAACCGGATCTGTCAGATGTGTGCGGCAAGAACCCAAAGTTG 1800
DB 2184 GACCGTGTCAAGAACTTAAACCGGCTGTGTATGTGTATGCGAAGAACTTCAAGTTGAG 2243
QY 1801 GAGTTGGCTAACTCGTGTGTGTATGATGTATGATGAGCAAGAACTTAAAGATTGGA 1860
DB 2244 GAATCTCGCAACTGT 2303
QY 1861 GAGAAAGCTGAATGAGAAAGAAATGTTTGAAGTATGACAACTTGAACCGCCAA 1920
DB 2304 GAGAGCTGTGATGATGAAGAAATGATGACCTCATGAAAGTACAGCTGATGTGGCAG 2363
QY 1921 TTCAGATGATATCATCTCAATGAAACGAATCCGAATGTTGAACCTTACCAGTACTT 1980
DB 2364 TTCAGATGATATCTCTCCAGATGAAACCGGATGAGAACTTACCGCTACATC 2423
QY 1981 TGCGACAGAAAGGAGCTTGTGTATGAGCTGATGTATGAGCTTGTGATGATGATGAT 2040
DB 2424 TGTCACAGAAAGGAGTCTTGTGTATGAGCTGATGTATGAGCTTGTGATGATGATG 2483
QY 2041 GTGAGGCAATGATCTTGCGGTTTGCACATTCGCAACCTGTAAACGATGACCAAGCC 2100
DB 2484 GTTGAAGCCATGATCTGTGTATGATGATGATGATGATGATGATGATGATGATGAT 2543
QY 2101 ATTATTTGTCATGGAATCTGTGTGTGAAATGATGATGATGATGATGATGATGATGAT 2160
DB 2544 ATCATTTGTCATGGAATCTGTGTGTGAAATGATGATGATGATGATGATGATGATGAT 2603
QY 2161 GACATCTCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
DB 2604 GAGCTTCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2663
QY 2221 TCCCAAGAGGCTTGAAGCAATGAGAGAGATATACATGAAAGATTTACTCGGAGAGA 2280
DB 2664 TCAAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2723
QY 2281 CTATTTGACCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
DB 2724 CTGTGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2783
QY 2341 GAGAGTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
DB 2784 GAGAGTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2843
QY 2401 GTTCCATTTGCAAGAGAGTAA 2422

DB 2844 GTTCTCCGCTGTGAGTAA 2865
RESULT 12
US-09-938-842A-1620
; Sequence 1620, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1620
; LENGTH: 2427
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1620
Query Match 56.6%; Score 1486.6; DB 9; Length 2427;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 1841; Conservative 0; Mismatches 570; Indels 3; Gaps 1;
QY 4 GCTGAGGTGCTCTCACTCGGCTCACAAGTCTCCGTGAGCGTTTGATGAGACCTTCTT 63
DB 10 GCTGAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 69
QY 64 GCTGAGGAAAGAGATTTTGGGCTTCTGCAAGATGAGAGGCAAGAAAGAAAT 123
DB 70 TCTGAGGAAAGAGATTTTGGGCTTCTGCAAGATGAGAGGCAAGAAAGAAAT 129
QY 124 CTGCAACATCAATTAATCTAGATTTGAAGCTATCTCTGAAGAGAAAGAAAG 183
DB 130 TTACAAGAAAGAGATTTTGGGCTTCTGCAAGATGAGAGGCAAGAAAGAAAG 189
QY 184 CTGCTATGATGATTTTGAAGATGATGATGATGATGATGATGATGATGATGATG 243
DB 190 CTGAGATGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 249
QY 244 CCATGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 303
DB 250 CCATGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 309
QY 304 GTTCAAGGCTTGTGTGTGAAGAGTCACTGTTGCTAGATCTCACTTCAAGAAAG 363
DB 310 CTGATGCTGTGTGTGAAGAGTCACTGTTGCTAGATCTCACTTCAAGAAAG 369
QY 370 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 429
DB 424 TCATCATTTCCCGGCAACCTTCAAAATCCATGATGATGATGATGATGATGATGATGAT 483
QY 430 GCGTCTATCCCTGTCACACCTCACAATATGATGATGATGATGATGATGATGATGATG 489
DB 484 GGTCACTTTCGGAAGATTTGTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 543
QY 490 GGTCACTTATTCGTAAGCTTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 549
DB 544 CTGAGATTCATTTGCAAGAGGCAAGAGATGATGATGATGATGATGATGATGATGATG 603

Db 550 CTTCGCTTCAACAGCCACGAGGCAAGAACTGATGTTGAGCGAAGANTTCAGAACTCTC 609
Qy 604 AATGCTCTTCAACATGTTTGAAGAAACAGAGAGATCTTGGTACCTCCTCTGAG 663
Db 610 AACCTCTGCAACACCTTGAAGAAACAGAGAGATCTTGGTACCTCCTCTGAG 669
Qy 664 AACCATGTCGCAATTCGAACACCGGTTCCAGAAATCGGTTTGAAGAGGTTGGGGT 723
Db 670 ACACTGATGAAAGTTTGAAGCCAGTTTGAAGAGATGTTGTTGAGAGGGATGGGGA 729
Qy 724 GACACCGCAAGAGCGTCTGAGATGATCCATCTCTTTGGATCTTCTTGAAGCAACT 783
Db 730 GACATGAGAGCGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
Qy 784 GATCTCTGACCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 843
Db 790 GATCTCTGACCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
Qy 844 CTCACTCCCAACGAGTACTTGGCTCAAGACATGTTTGGGGTATCCCGACACCGGTTGG 903
Db 850 CTCTCTCAAGATGTTTACTTGGCTCAAGACATGTTTGGTTACCTGACACTGGTGGGA 909
Qy 904 CAGGTTGTTACATCTTGAATCAAGTCCGAGCTTTGAAGATGAGATGATGATGATGAT 963
Db 910 CAGGTTGTTACATCTTGAATCAAGTCCGAGCTTTGAAGATGAGATGATGATGATGAT 969
Qy 964 AAGCAACAGAGCTCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1023
Db 970 AAGCAACAGAGCTCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1029
Qy 1024 GCTGTCGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1083
Db 1030 GCGGTAGAACTACATGCGGTGAAGCTCTGAGAGAGTTTATGATGATGATGATGAT 1089
Qy 1084 ATTCTTGAAGTCTCTTCAAGACAGAAAGGATGTTTCAAAATGATCTCAAGATT 1143
Db 1090 ATTCTTGAAGTCTCTTCAAGACAGAAAGGATGTTTCAAAATGATCTCAAGATT 1149
Qy 1144 GAAAAAGTCTCTTCAAGACAGAAAGGATGTTTCAAAATGATCTCAAGATT 1203
Db 1150 G--AAGTCTCTTCAAGACAGAAAGGATGTTTCAAAATGATCTCAAGATT 1206
Qy 1204 GAGTTGACGAGCGGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1263
Db 1207 GAATTTGATGAGACCTGATCTTATCTGATGATGATGATGATGATGATGATGATGAT 1266
Qy 1264 TCTTGTCTGACATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1323
Db 1267 TCTTGTCTGACATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1326
Qy 1324 ACAAATATCCAGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1383
Db 1327 ACAAATATCCAGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1386
Qy 1384 TGCCAAATTAAGATGATCTTCTTGAAGACATGATGATGATGATGATGATGATGAT 1443
Db 1387 TGCCAAATTAAGATGATCTTCTTGAAGACATGATGATGATGATGATGATGATGAT 1446
Qy 1444 TTCCAGAAATTAAGATGATCTTCTTGAAGACATGATGATGATGATGATGATGAT 1503
Db 1447 TTCCAGAAATTAAGATGATCTTCTTGAAGACATGATGATGATGATGATGATGAT 1506
Qy 1504 ACTCTTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1563
Db 1507 ACTCTTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1566
Qy 1564 ATTGTTTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1623
Db 1567 ATTGTTTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1626
Qy 1624 TTGAAGATTTTCACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1683
Db 1627 TTGAAGATTTTCACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1686

Qy 1684 CACTTATGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1743
Db 1687 CACTTATGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1746
Qy 1744 CGTGTCAGAACTTAAACCGGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1803
Db 1747 CGTGTCAGAACTTAAACCGGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1806
Qy 1804 TTGCTCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1863
Db 1807 CTAGCTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1866
Qy 1864 AAGCTGAAATGAAGAAATGTTGATGATGATGATGATGATGATGATGATGATGAT 1923
Db 1867 AAGCTGAAATGAAGAAATGTTGATGATGATGATGATGATGATGATGATGATGAT 1926
Qy 1924 AGATGATATCATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1983
Db 1927 AGATGATATCTCTCTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1986
Qy 1984 GACAGAAAGGATCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2043
Db 1987 GACAGAAAGGATCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2046
Qy 2044 GAGGCAATGATCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2103
Db 2047 GAGGCAATGATCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2106
Qy 2104 ATTGTCATGAGAAATCTGTTTCAACATGATGATCTTCAACATGATGATCTTCAAC 2163
Db 2107 ATTGTCATGAGAAATCTGTTTCAACATGATGATCTTCAACATGATGATCTTCAAC 2166
Qy 2164 ATACTGTCATCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2223
Db 2167 ATACTGTCATCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2226
Qy 2224 CAAGGAGCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2283
Db 2227 CAAGGAGCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2286
Qy 2284 TTGACCTGACAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 2343
Db 2287 TTGACCTGACAGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 2346
Qy 2344 AGTGTGTTTACCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2403
Db 2347 AGTGTGTTTACCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2406
Qy 2404 CCATTGACAGAGA 2417
Db 2407 CCTCTGACAGAGA 2420

RESULT 13
US-09-938-842A-1620
; Sequence 1620. Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111

; PRIOR FILLING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1620
; LENGTH: 2427
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1620

Query Match 56.6%; Score 1486.6; DB 11; Length 2427;
Best Local Similarity 76.3%; Pred. No. 0;
Matches 1841; Conservative 0; Mismatches 570; Indels 3; Gaps 1;

Qy 4 GCTGAGCGTGTCTCATCTCGCGTCCACAGTCTCCGTGAGCGTTGGATGAGACCTCTCTT 63
Db 10 GCTGAGCGTATGATATACCGCGCTCCACAGCCGATGAGCGTTGAGAGAACTCTGT 69
Qy 64 GCTCAGAGGAGAGATTTTGGCCCTTGGCTCAGAGATCGAGGGCAAGAAAGAAAT 123
Db 70 TCTGAGAGAAAGAAAGTCTTGGCTTTCAGAGGTTGAAAGCCAAAGTAAAGTAT 129
Qy 124 CTGCAACACATCAAAATTATCTAGAGTTGAAAGCTATCCCTGAAAGAGAAAGAAAG 183
Db 130 TTACAAACAAACAGATCATTTGCTGAATTCGAAAGCTTTCCTGAACAAACCGAGAA 189
Qy 184 CTCGCTAATGATGATTTTGGAAATTTGAAGCTAGTCAAGAAAGCATGCTGTGCT 243
Db 190 CTGGAAGGTGCTCTTCTTGAACCTCTCAATCTCAGAGCAATTTGTGCTCA 249
Qy 244 CCAAGGTTGATCTTGTCTGTCAGAGCGCTGTTGGAGAGTACATTAAGATGAT 303
Db 250 CCAAGGTTGCTCTAGCTGTAAGGCAAGGCGTGTGTTGGAAATCTTACAGATCAT 309
Qy 304 GTTCAAGCGCTTGTGTGAGAGAACTCATGTGCTGATCTTCACTTCAAGAAAGAG 363
Db 310 CTCATGCTCTTGTGTTGAAGAACTCAACCTGCTGATTTCTTCAATTCAGAGAA 369
Qy 364 CTGTGATGAGAGTTCAATGAGAACTTTGTTGATTTGAAATTTGAGCCCTTCAAC 423
Db 370 CTGTGATGAGAGTTAAGATGATTAATTTCACTTGTAGCTTGAATTTGAGCCATTCAT 429
Qy 424 TCATCATTTCCCGCCCACTCTTCAAAATCCANTGGTAAATGGTGGAGATTCCTAAT 483
Db 430 GCGTCAATCCCTCGTCAACACTCACAAAATCATTTGAAATGGTGTGACCTTCTTAC 489
Qy 484 CGTCACTTTGCGCAAAATTTGTCATGACAAAGAGACATGCAACCTTGTCTGAATTC 543
Db 490 CGTCAATTTATGGCTTAAGCTCTTCCATGACAAAGAGATTTGCTTCCATTCCTTAAGTTC 549
Qy 544 CTCAAGATCTATGTCACAAAGGCAAGAACTGATGTTGAATGACAGAAATTCAGAACTTG 603
Db 550 CTTCGTCTTCAAGCCACCAAGGCAAGAACTGATGTTGAGCGAGAAATTCAGAACTTC 609
Qy 604 AATGCTCTTCAACATGTTTGAAGAAAGCAAGAGATCTTGTGATCCCTACCTCCAG 663
Db 610 AACACTTGCACACACCTTGAAGAAAGCAAGAGATCTTGAAGAGCTTAAAGTCCGAA 669
Qy 664 AACCATGTGCGCAATTCGAAACACCGTTCAGAGAAATCGGTTTGAAGAGTGGGCT 723
Db 670 ACACTGATGAAGAGTTTGAAGCCAAAGTTTGAAGAGATTTGCTTGAAGAGGAGATGGGA 729
Qy 724 GACACCGAGAAACGCGTCTGAGATGATCAACTCTTTTGGATCTTCTTGAAGCAACT 783
Db 730 GACAAATGAGAGCGGTCTTGAACATGATGATCTTCTTGGACCTTCTTGAAGCGCT 789
Qy 784 GATCTTGGACCTTGGAGAAAGTCTTGGAGAAATCCCATAGGTCTTCAATGTTGTGAT 843
Db 790 GATCTTGGACCTTGGAGAAAGTCTTGGAGAAATCCCATAGGTCTTCAATGTTGTGAT 849
Qy 844 CTCATCTCCCAAGGATCTTGGCTCAAGACATGTTTGGGATATCCGACACGAGTGG 903
Db 850 CTCATCTCAACATGATTTACTTGGCTCAAGACATGTTTGGATATCCCTGACACATGATGGA 909
Qy 904 CAGGTTGTTTACATCTTGGATCAAGTCCGAGCTTGGAGAAATGAGATGCTCTCGTATA 963

Db 910 CAGGTTGTTTACATCTTGGATCAAGTCCGCTCGAGATGAGATGAGCTTCAACGTAAT 969
Qy 964 AAGCAACAAGAACTCAACATCAACCCCTGATCCCTCATTAATTAAGATCTTCTCGAT 1023
Db 970 AAGCAACAAGAACTCAACATTAATAACCAAGATTTCTATTCTAAGCTTCACTGAT 1029
Qy 1024 GCTGTGGAACCAATGCGGTCAACGACTTGAAGAAATATACGAAACAGAGACTCGAT 1083
Db 1030 GCGGTAGAACTACATCGGTGAACGCTCGAGAGATTAATGATTTCTGAGTACTGTGAT 1089
Qy 1084 ATTTTGGATACCTTCAAGAACAGAAATTTGTTGAAATTTGATTTCTCAAGTTT 1143
Db 1090 ATTTTGTGTGCTTCAAGAACAGAAAGGATTTGTTGCAAAATGATCTCAAGTTC 1149
Qy 1144 GAAAGATCTGGCACTTCTGAAAACCTTACACAGAGATGTTGCTCATGAATCTCAAA 1203
Db 1150 G--AAGTCTGGCACTTCTGAGACTTACACAGAGATGCTGGGTTGAGCTATCGAA 1206
Qy 1204 GAGTTGACCGCAACGCGATCTGATCATCGAAACCAACGCGCAATATCGTGGC 1263
Db 1207 GAATTTGAATGGCAAGCTGACCTTATCATTTGATTAATCAAGTATGGAATCTTGTGCT 1266
Qy 1264 TCTTGTCTGCACTAATTAAGTGTCAACAGTGCACATCGCCCATGCTTGGAGAAAG 1323
Db 1267 TCTTATTTGGCTCAAACTTGTGTCACTCAGTATCAATGTCTATGCTTGAAGAAA 1326
Qy 1324 ACAAAATATCCAGATTCAGATATCTATTTGGAAGAGCTTGAAGCAATATCATTTCTCT 1383
Db 1327 ACBAAGTACCGGATTTGTGATTTCTACTGGAAGAGCTTGAAGCAATATCTTCTCA 1386
Qy 1384 TGCCAAATTAACGTGATCTTTTTCGAATGAACCAATCAAGATTTCAATCACCAGTACT 1443
Db 1387 TGCCAGTTCACTGGGATATTTTTCGAATGAACCACTGATTTCAATCACAATGATCT 1446
Qy 1444 TTCCAGAAATTTGAGAGAACCAAGACATGTTGTCAATACAGAGCCCACTGCTTTC 1583
Db 1447 TTCCAAAAATTTGTGAGAACCAAGAACTGTTGGGAGATGAAACCAACACAGCTTT 1506
Qy 1504 ACTCTTCTGATCTTCAACCGTGTGTATCATGATGATGATGATTTCCCAATTCAC 1563
Db 1507 ACTTCTCCGATTTGTATCCAGTTTTCACGGATTTGATGATTTCCCAATTCAC 1566
Qy 1564 ATTGTTTCCCTGTGTGATGATGATGATTAATCTTCTTCAACAGAGAGAGCCGAGG 1623
Db 1567 ATTGTTCTCTGTGTGATGATGATGATTAATCTTCTTCAACAGAGAGAGCCGATGA 1626
Qy 1624 TTGAAGATTTCCATCTGAGATGAGAACTTTTTCACCAAGTGTGAATGAAGAA 1683
Db 1627 TTGAATAGTTCCACTTGAATGAGAGAGCTCTTCAACGCGATGTTGAAGAAAGAG 1686
Qy 1684 CACTTATGTGTGCTCAATGACCGCAACAGCCAAATCTGTTCACAATGSCCAAGCTTGAT 1743
Db 1687 CACTTATGTGTGCTCAAGACAGAGACAGAAAGCCCAATCTTCAACATGSGCTGAT 1746
Qy 1744 CGTGTCAAGAACTTAAACCGGACTCTGTGAGTGTGTGCGGCAAGAAACCAAGTTGCTGAG 1803
Db 1747 CGTGTCAAGAACTTGTGAGGCTTGTGAGTGTGATGAGGAAAGAACACCGCTTGTGAG 1806
Qy 1804 TTGCTTAACCTCGATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1863
Db 1807 CTAGCTTAATGTTGTTGTTGAGAGAGACAGAGGAAAGTCAAAAGATCAATGAAGAG 1866
Qy 1864 AAGCTGAATTAAGAAATGTTTGAAGCTGATGACAGATGACAACTTGAACGCGCAATTC 1923
Db 1867 AAGCTGAATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1926
Qy 1924 AGATGATATCATCTCAATTAAGACAGATCGAAATGTTGAATCTTACCGATACATTCG 1983
Db 1927 AGTGTATCTCTCTGATGATGAGACCGGTTAAGAAACGTTGAGCTGACCGTATCATCTGT 1986
Qy 1984 GACAGAAAGTGTGCTTGTGACAGCTGCAATTTGATGAAGCTTTGGATTTGACAGTTGTG 2043

Db 1987 GACACCAAGGAGCTTTTGTCCAACTGCAATTAATGAAAGCTTTGGTTAACTGTTGTG 2046
Qy 2044 GAGGCAATGACTGTGGGTTTGCCAACTTGGCAACTGTGAACGGTGAACGCGAGATT 2103
Db 2047 GAGGCTAATGACTGTGGTTTACCGACTTTCGCACTTCAAGAGGTGTCAGCTGAGATC 2106
Qy 2104 ATTGTGCAATGGAATCTGTTTCAACATTTGATCTTCAATGATGATCAAGCTGCTAC 2163
Db 2107 ATTGTGCAATGGAATCTGTTTCAACATTTGATCTTCAATGATGATCAAGCTGCTAC 2166
Qy 2164 ATACTGCTGATTTCTTTGAAAAGTGAAGAAAAGTCAATCTCACTGGATTAAGTCTCC 2223
Db 2167 ACTGTGCTGATTTCTTTCAACAGTGAAGAGATTCATCTCACTGGATTAAGTCTCA 2226
Qy 2224 CAAGGAGCTTGAACGAATGAGAGAGATTAATGAGAAATTACTCGAGAGACTA 2283
Db 2227 AAAGGAGGCTTCAAGAGATTAAGAGAAATACCTTGGCAATCTATTCACAGAGCTC 2286
Qy 2284 TTGACCTGACAGAGATGATGATTTGGAAGCATGTTTCAACCTTGAACGCGTGAG 2343
Db 2287 TTGACATTTGACTGTGTGTATGATTTCTGGAAGCATGTCTGAACTTGAACCTTTGAG 2346
Qy 2344 AGTGTGCTTACCTTGAATGTTTATGCTCTTAAGTACCTTGAAGCTGAGTCACTT 2403
Db 2347 GCTGCGCTTACCTTGAATGTTTATGCTCTTAAGTACCTTGAAGCTGAGTCACTT 2406
Qy 2404 CCATTGGCAGAGA 2417
Db 2407 CCTCTTGACAGA 2420

RESULT 14

US-10-425-114-9355
Sequence 9355, Application US/10425114
Publication No. US2004003488A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT FILING DATE: 2003-04-28
CURRENT APPLICATION NUMBER: US/10/425.114
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 9355
LENGTH: 2279
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 700843719_FLI
US-10-425-114-9355

Query Match 50.7%; Score 1331; DB 13; Length 2279;
Best Local Similarity 78.8%; Pred. No. 0;
Matches 1599; Conservative 0; Mismatches 426; Indels 3; Gaps 1;

Qy 404 TGAATTTGAGCCCTTCAATCATATCCCGGCGCAACTTTCAAAATCCATTGTA 463
Db 1 TGAATTTGAGCCCTTCAATCATATCCCGGCGCAACTTTCAAAATCCATTGTA 60
Qy 464 ATGTGTGAGATCTCTTAATGTCACCTTTGGCAAAATGTTCCATGACAGAGACA 523
Db 61 ATGTGTGAGATCTCTTAATGTCACCTTTGGCAAAATGTTCCATGACAGAGACT 120
Qy 524 TGCACCTTTGCTGCAATTTCTCAAGTTCATTTGACAGAGGCAAGAACTGATGTA 583
Db 121 TGCACCTTTGCTGCAATTTCTCAAGTTCATTTGACAGAGGCAAGAACTGATGTA 180
Qy 584 ATGACAGATTCAGAACTTGAATGCTTTCAACATGTTTGAAGAAAGAGAGATATC 643

Db 181 ATGACAGATTCAGAACTTGAATGCTTTCAACATGTTTGAAGAAAGAGATATC 240
Qy 644 TTGTGACCTTACCTCTGAGACACCATGTGCCAAATTTGAACACCGGTTCCAGAGAAATCG 703
Db 241 TGGGACAGTGTCTCTGAAATCTCCCTCACTGAAATTTGACACAAAGTTCCAGAGATTTG 300
Qy 704 GTTTGAAAGAGTTGGGGTGACACCGCAGAACCGTGTGAGATGATTCACCTCTT 763
Db 301 GTTTGAAAGAGTTGGGGTGACACCGCAGAACCGTGTGAGATGATTCACCTCTT 360
Qy 764 TGAATCTTCTTGAAGCACTGATCTTGCACCTTGAAGATTTCTTGGAGAAATCCCA 823
Db 361 TGAATCTTCTTGAAGCACTGATCTTGCACCTTGAAGATTTCTTGGAGAAATCCCA 420
Qy 824 TGTGTTCATGTTGTGATTTCTCACTCCCAACGATTTCTGCTCAAGACAATGTTTG 883
Db 421 TGTGTTCATGTTGTGATTTCTCACTCCCAACGATTTCTGCTCAAGACAATGTTTG 480
Qy 884 GGTATCCCGACACCGGTGGGAGTTGTTATCTTGTGATCAAGTCCGAGCTTTGAGA 943
Db 481 GATACCTTGACATGTGTGACCAAGTTGTTACATCTTGAATAGTTGTGCTTTGAGA 540
Qy 944 ATGAGATCTCTCTGATTAAGCAACAGACTCAACATCAACCTTGAATCTCTATTA 1003
Db 541 ACGAGATCTCTGATTAAGCAACAGACTCAACATCAACCTTGAATCTCTATTA 600
Qy 1004 TTACTAGACTTCTCTGATGATCTGTGCAACACATGCGGTCAACGACTTGAAGAAAT 1063
Db 601 TCACCCGTCTCTCTGATGATCTGTGCAACACATGCGGTCAACGACTTGAAGAAAT 660
Qy 1064 ACGGACAGACACTCGGATTTCTTGTGATGATCTTGAACAGAAAGGAAATGTTTC 1123
Db 661 TCGAACCGAGACACTCGGATTTCTTGTGATGATCTTGAACAGAAAGGAAATGTTTC 720
Qy 1124 GAAATGATCTCAAGATTTGAAAAGTCTGCGCATTACTTGGAAACCTTACAGAGATG 1183
Db 721 GTGAGTGTATCTCAAGATTTG---AAGTCTGCGCATTACTTGGAAACCTTACAGAGATG 777
Qy 1184 TTGCTCATGAATCTCCAAAGATTTGACGCGCAACGCGCATGATCTATTCGAAACNCA 1243
Db 778 TTGCTCATGAATCTCCAAAGATTTGACGCGCAACGCGCATGATCTATTCGAAACNCA 837
Qy 1244 GCGACGCAATATGTGCGCTCTTGTGCACTTAATTTAGGTGTCACAGTGCACA 1303
Db 838 GTGATGAAACATTTGTGCTCTTGTGCACTTAATTTAGGTGTCACAGTGCACA 897
Qy 1304 TCGCCATGCTTTGAGAGAGCAAAATATCCAGATTCAGATATCTATTTGAAGAGCTTG 1363
Db 898 TTGCTCATGACCTTGAAGAGCAAAATATCCCGAATCCGACATTTACGAAAAATTTGG 957
Qy 1364 AAGACAAATACATTTCTCTTGCATTAATTAAGCTGATCTTTTGAATGAACATACAG 1423
Db 958 AAGACAAATACATTTCTCTTGCATTAATTAAGCTGATCTTTTGAATGAACATACAG 1017
Qy 1424 ATTTCATATACAGATCTTTCAGAGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAG 1483
Db 1018 ATTTCATATACAGATCTTTCAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAG 1077
Qy 1484 ACGAGAGCACTGCTTCACTCTCTGCTCTCAACCGTGTGATCATGATGATGAT 1543
Db 1078 ATGAGTCTCAACAGCTTTTACCTTCTGAGACTTCAACCGTGTGATCATGATGATG 1137
Qy 1544 TGTGTATCCCAATTCACATTTGTTCCCTGTGCTGATATGAGATTAATCTTCCCTT 1603
Db 1138 TGTGTATCCCAATTCACATTTGTTCCCTGTGAGATTAATTAATCTTCCCTT 1197
Qy 1604 ACACGAG 1663
Db 1198 CACCGAG 1257
Qy 1664 CCAAGTTGAGATGAG 1723

```
Db 1258 GCTCTGTGAGAAATGAAACACATATGCGTGTGAGAGACCGCAGCAAGCCGATTAATCT 1317
Qy 1724 TCACAATGCCAAGGCTTGATCGTGTCAAGAACTTAACCGGACTCTGTGAGTGTGGCGCA 1783
Db 1318 TCACCATGGCAAGGTTGGACCGGTGTGAAGAACATCACAGGACTGTGTGAGTGTGACGGTA 1377
Qy 1784 AGAACCCAAAGTGGTGTGAGTGTGCTTAACCTGTGATGTGTAGTGTGTATAGCGCAAGG 1843
Db 1378 AGAACCGCAACCTGGAGGAGTGTGTGAACCTTGTGTGTGTGTGCGCGAAGACGAGAGAGG 1437
Qy 1844 AATCTAAAGATTTGGAGAGAGAGGCTGAATGAAAGAAATGTTTGAAGTGTGACCAAGT 1903
Db 1438 AGTCAGAGACTTGGAGAGAGAGGCGAGATGAGAGATGATGAGCGCTGTGACAGACCT 1497
Qy 1904 ACAACTTGAACGGCCCAATTCAGATGATATCATCTCAATGAAACGAATCCGAAATGTTG 1963
Db 1498 ACAAGTTGAACGGGCAATTCAGATGATTTATCTCAGATGAAACGCTGTGAGGAACGGAG 1557
Qy 1964 AACTTTACCGATCACTTTGGCAGACAGAAAGTGTCTTGTGTACAGCTGTGATTTGTAG 2023
Db 1558 AGCTGTACCGTGTGATCTGCGCACACCAAGGAGCTTGTGTGACGCGCGCTATATACGAG 1617
Qy 2024 CTTTGTGATTTGACGTGTGGAGGCAATGACTTGGGTTTGGCAACATTTGCAACCTGTA 2083
Db 1618 CTTTGTGTGTGACATGTGTGTGAGGCAATGACTTGTGTGTGCAACATTTGCGCACATGCA 1677
Qy 2084 ACGGTGACCAAGCCGAGATTAATGTTCATGAGGAAATCTGTTCACATGATCTCTTACC 2143
Db 1678 ATGTGTGCTGTGTGATGATCATTTGTGACATGAGCAATCTGGCTTCAATGACCTTACC 1737
Qy 2144 ATGTGTATCAAGCTGTGTGATCATCTGTGTGATTTCTTTGAAAGTGTAAAGAAATCAT 2203
Db 1738 ATGTGTACCGGTGTGTGTGTATCTCTTGTGTGACTTCTTTGAGAAAGTGAAGCTTGAACCA 1797
Qy 2204 CTCACGTGGATTAAGATCTCCCAAGAGGCTTGAAGCAATGAGAGAAATGATATACATGGA 2263
Db 1798 CCCACTGGAGAAACATCTCAAGAGGCTGTGTCTCCAGCGATTTGAGAGAAATGATACATGAC 1857
Qy 2264 AGATTACTCGGAGAGACTATTTGACCTGTGACAGAGATGATGATTTCTGGAACATGTTT 2323
Db 1858 AAATTTTATCAACAGAGGCTTGTCACTGTCACTGTGTGTATGAGCTTCTGGAAGCATGTGT 1917
Qy 2324 CCAACTTTGAAGCGCTGTGAGAGTGTGTGTACCTTGAATGTTTATGCTCTTAAGTACC 2383
Db 1918 CTAACCTTTGACCGCGGTGAGAGCCCGCTATCTCGAGATGTTCTATCTCTCAAGTACC 1977
Qy 2384 GTAAGCTGTGATCAATCAATTCATTTGAGAGAGAGATTAATGAACCTG 2431
Db 1978 GCAAATTTGACCGAGTCTGTGTGCTGTGTGTGATTAACAGAGATG 2025

RESULT 15
US-10-289-757-133
; Sequence 133, Application US/10289757
; Publication No. US20030180751A1
; GENERAL INFORMATION:
; APPLICANT: Demmer, Jeroen
; APPLICANT: Forester, Richard L
; APPLICANT: Gibson, John Bryan
; APPLICANT: Shenk, Michael Andrew
; APPLICANT: No. US20030180751A11189, Geoffrey
; APPLICANT: Glenn, Matthew
; APPLICANT: Saulsbury, Keith Martin
; APPLICANT: Hall, Claire
; TITLE OF INVENTION: Compositions isolated from forage
; TITLE OF INVENTION: grasses and methods for their use
; FILE REFERENCE: 11000.1061U
; CURRENT APPLICATION NUMBER: US/10/289,757
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/337,703
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 218
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 133
; LENGTH: 2714
; TYPE: DNA
; ORGANISM: Lolium perenne
US-10-289-757-133

Query Match 49.7%; Score 1305.4; DB 15; Length 2714;
Best Local Similarity 72.1%; Pred. No. 0;
Matches 1731; Conservative 0; Mismatches 662; Indels 9; Gaps 2;

16 CTCACGTGGGCGCCACAGTCTCCGTGAGGCTTGGATGAGACCTTCTGTGACAGAAC 75
57 CTGACTGTCTCCACAGTCTCCGAGAGCGCTTGTGTGACCTTCTCTCCATCTTAC 116
76 GAGATTTTGGCTTGTCTCAAGATGACAGGCGCAAGGAAAGAAATTTCTCAACACCAT 135
117 GAGCTCATTTGGCCCTTTTCAAGATTAATGTTTCAACAGGCGCAAGAAATGCTTCAAGGCGCAC 176
136 CAAATTTATCTAGAGTTTGAAGCTATCCCTGAAAGAAACAGAAAGAAAGCTGCTAATGT 195
177 CAGCTGTCTACAGTGTGAAAGCTTGTGTGAGGCTGCAAGAGAGA-----GATATGCA 230
196 GCATTTTGTGAAGATTTGAAGGCTAGTCAAGGAGCGATCGTGTGCTCCATGAGGTTGCA 255
231 CCTTTGAAGACATTTCTCCGTGTCTGAGAGCAATTTGTGTGCTGCCCATGAGGTTGCA 290
256 CTGTCTGTGTCTCCAGGCGCTGTGTGTGTGAGTGTATGAGTGAATGTTCAAGCCCTT 315
291 CTGGCATCATAGGCGCTAGGCGGTGTGTGTGAGTGTATGAGTGAATGTTAGCGAGTTG 350
316 GTTGTGAGAACTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 375
351 GCTGTGAGAGGCTGAGCAATTTCCAGATGATCTGTGATTTCAAGAAACAGCTTGTGTGTAG 410
376 AGTTCAATGAGAACTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 435
411 CAGCGCAGAGCAAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 470
436 CGCCCAACTCTTTGAAATCCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 495
471 CGTCTTCATATGTCATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 530
496 GCAAATTTGTTCATGACAGAGAGATGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 555
531 TCAGAGCTGTTCAGAGCAAGAGAGGCTTCAACCATCATGAACTTCTTAAGAGCCCAT 590
556 TGTCAAGAGGCGCAAGACATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 615
591 AACCAAGAGGCGACAGCAATGATCTTAAGAGCCGAATTCAGAGCTTGTGTGTGTGTGT 650
616 CAGTGTGTGAGAAAGAGAGAGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 675
651 TCAGCCCTTGAAGAGGAGAGAGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 710
676 GAATTCAGAACCGGTTCCAGAGAAATGAGTTGTGAAGAGTGTGTGTGTGTGTGTGTGTGT 735
711 GAGTTCAACCAAGAGTTCCAGAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 770
736 CGGCTGTGAGATGATCAATCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 795
771 GGTGTACAGACACATTCATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 830
796 CTTGAGAAATTTCTTGGAGAGATCCCATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 855
831 TTGAGAAAGTTCTTTGAACTATACATATGTTTCAATGTTGTATTCGTGTGTGTGTGTGT 890
856 GGATATCTTGTCTCAAGACATGTTTGGGATATCCGACACCGGTGGCAGGTTGTGTGT 915
891 GGATATTTGTCTCAATCAATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 950
916 ATCTTGATCAAGGCGAGCTTGTGAGAAATGAGATGTGTGTGTGTGTGTGTGTGTGTGTGT 975
951 ATCTTGATCAAGTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1010
```

Oy 976 CTCAATCACCCTCGAATCTCATTTACTAGATCTTCTGATGCTGCGAACA 1035
 Db 1011 CTTGACATACCCCTTAAGATCTCTATGTGCACAGGCTGTGCTGATGCTGTGAACT 1070
 Oy 1036 ACATGCGGTCAACGACTTGAGAAAGTATACGGAACAGACACTCGAATATCTTGAGTA 1095
 Db 1071 ACATGTGCGCAGCGCTGAGAAAGTTATGTGGATCTGAGCACACTGACATCTCCGTGT 1130
 Oy 1096 CCGTTCAAGACAGAAAGGGAATTTGTGAAAATGATCTCAAGTTTGAAAAATCTGCG 1155
 Db 1131 CCATTCAACCTGAGAGGGGATCTCCGTAGTGAATCTCTCGTTTG--ATGTCGCG 1187
 Oy 1156 CCATCTCTGGAACCTACACAGAGATGTGCTCATGAAATCTCAAGAGTTGACAGCGC 1215
 Db 1188 CCAATCTCGAGACATACACGAGATGTTGCAACGAACTCATAGAGAAATGACAGAC 1247
 Oy 1216 ACGCCAGATCTGATCATCGAAACNACAGCGCAATATCGTGCCTCTTGCTGCA 1275
 Db 1248 AAGCTGATCTCATATTGGCAACTACAGTACGGTTAACCTGTGCGCACTGCTTGG 1307
 Oy 1276 CATTAATTAAGTGTACACAGTGCACATCGCCCATGCTTTGAGAAAGAAATATCA 1335
 Db 1308 CATAGTTGGAGTTACCCAGTGCACATGCCCATTGCTTGAGAAAACAAAATACCC 1367
 Oy 1336 GATTGAGATCTATTTGGAAGAGCTTGAAAGCAAAATCAATTTCTTGGCAATTTCA 1395
 Db 1368 AACTGACATATATTTGACAAATTCAGACGCAATTCATTTCTCATGCGCACTGCT 1427
 Oy 1396 GCTGATCTTTTGGCAATGACCATACAGATTCATCATCACCAATCTTCCAGGAAAT 1455
 Db 1428 GCTGACCTTATTTGCCATGATACACTGATTTCACTACACAGACATTCAGGAAAT 1487
 Oy 1456 GAGAGAGCAAGAGACACTGTTGCTCAATACAGAGCCACACTGCTTTCACTTCTG 1515
 Db 1488 GCTGAGACCAAGAGAGCGGTGGCCAGATGAGTCTCAATGCTTCACTCCCTGAT 1547
 Oy 1516 CTCTACCGTGTGTACATGATGATGATGTTGATCCCAATTCACATGTTTCCCT 1575
 Db 1548 CTCTACCGTGTGTACATGATGATGATGATGATGATGATGATGATGATGATGAT 1607
 Oy 1576 GGTGCTGATATGAGATATACTTCCCTTACACCGAAGAGAGCGGATGGAAGCATTC 1635
 Db 1608 GAGAGAGATATGAGTGTCTACTTCCCATACCTGAACCTGACAAAGGCTTACTGCTTC 1667
 Oy 1636 CATCTGATGATGAGAACCTTTTACACCAAGTTGAGATGAAAGACCTTATGTGTG 1695
 Db 1668 CACCTGAAATTTGAAGACTCTTTTACAGTGAATGTTGAGAACTTGAAACAAAGTTGTA 1727
 Oy 1696 CTCAATGACCGCAACAAGCCAAATTTGTTCACAATGCCAAGGCTGATCGTGTCAAGAC 1755
 Db 1728 TTGAAGACAAAGAACGCTATCATCTTCAATGCTCGTGTGACCGTGTGAAGAC 1787
 Oy 1756 TTAACCGGACTCGTGAAGTGTGCGGCAAGAACCCAAAGTTGCGTGAAGTCACTC 1815
 Db 1788 ATGACAGGCTGTGATGATGTTGTGTAAGATGCTCATCTGAAGATTTGGCAACCTT 1847
 Oy 1816 GTAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1875
 Db 1848 GTGATGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1907
 Oy 1876 AAGAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1935
 Db 1908 AAGAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1967
 Oy 1936 TCTCAATGAAACAGATCCGAATGTTGAATTTTACCGATACATTTGCGACACGAAAGT 1995
 Db 1968 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2027
 Oy 1996 GCGTTTGTACAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2055
 Db 2028 GCATTTGTTACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2087

Oy 2056 TCGGTTTGCCCAATTCGCAACCTGTACAGGTGAGACCGCAGAGATTATGTCATGGG 2115
 Db 2088 TGTGTTTGCCCAATTCGCAACCTGTACAGGTGAGACCGCAGAGATTATGTCATGGG 2147
 Oy 2116 AAATGTGTTTCAACATTTGATGCTTACATGATGATGATGATGATGATGATGATGAT 2175
 Db 2148 GTGCTGTGTTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2207
 Oy 2176 TTTCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2235
 Db 2208 TTTCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2267
 Oy 2236 AAACGATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2295
 Db 2268 AAGAGATTTATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2327
 Oy 2296 GAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2355
 Db 2328 GTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2387
 Oy 2356 CTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2415
 Db 2388 CTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2447
 Oy 2416 GA 2417
 Db 2448 GA 2449

Search completed: June 5, 2004, 09:02:07
 Job time : 1545 secs

THIS PAGE BLANK (USPTO)